

Soft responses to misconduct

The Federation of American Scientists for Experimental Biology and the Association of American Medical Colleges lead the 'heads-in-the-sand' school on the scientific misconduct issue.

The Office of Research Integrity (ORI) at the US health department continues to plough a lonely furrow as it tries to tackle misconduct. Each investigation that it mounts is complex and problematic for everybody involved (see *Nature* **419**, 332–333; 2002). Its administrators have sensibly concluded that its mission will be best accomplished not through investigation and punishment, but by spreading awareness of the importance of ethical conduct throughout the US life-sciences community. Sadly, its efforts are not receiving the support they deserve from US scientific societies and medical schools.

Most recently, the ORI has proposed a survey of National Institutes of Health grant applicants to determine perceptions of various aspects of scientific conduct. Government regulations require the office to publish its proposed set of questions in the Federal Register for comments from interested parties. True to form, scientific societies have lambasted it for allegedly meddling in areas beyond its purview.

In particular, the Federation of American Societies for Experimental Biology (FASEB) and the American Association of Medical Colleges (AAMC) have attacked the survey (see www.faseb.org) for daring to seek out information on the pervasiveness of low-key unethical behaviour, such as authors citing papers that they haven't read. FASEB and the AAMC say they are indignant that the ORI is seeking to measure misbehaviour that falls outside the scope of the tight 'fabrication, falsification and plagiarism' definition of scientific misconduct that the US federal government adopted a couple of years ago, after several years of low-octane wrangling.

But in a country wracked by allegations of unethical behaviour in business, politics and even in science, perhaps FASEB and the AAMC protesteth too much. For a start, the restrictive definition of misconduct that was accepted by the government represented a lowest-common-denominator approach, hammered out behind closed doors in rooms where FASEB and the AAMC may have been represented but other parties — notably patients' groups, and the taxpayer who foots the bill for the research — were notable only for their absence. The six-year process that derived the definition managed to take the

more radical findings of a commission chaired by the late Kenneth Ryan — which called for a much broader definition of misconduct — and squeeze it back into a bureaucratic strait-jacket wherein routine misbehaviour would go unremarked and unpunished.

The shredding of Ryan's tough prescription made sense to some research leaders at the time but, given the subsequent explosion of allegations of corruption in many quarters of US public life, one has to wonder if it still does. This weekend, for example, the country has been wondering whether to laugh or cry at news that an influential stock-market analyst once issued a 'buy' recommendation for AT&T in an effort to secure the admission of his two-year-old twins into an exclusive Manhattan nursery school. In this climate, scientific leaders would do well to work for transparency in science and for a strong and confident regulator in the ORI. Instead, they seem to want to thwart the watchdog agency at every turn.

The US life sciences have recently benefited from a major expansion in public funding, and have experienced relatively few instances of high-profile fraud. (In an unexpected turn, the most important recent cases of fraud have been in physics, which once considered itself more or less immune to the malaise.) The ORI is politically weak and may appear vulnerable to scientific societies and university administrators who have traditionally viewed it with suspicion.

By keeping track of researchers' perceptions on the need for vigilance with regard to misconduct, the ORI can play a useful role in providing information on standards of research conduct. But by attacking its every move, including efforts to introduce ethics training for graduate students, FASEB and the AAMC give a good impersonation of aged, out-of-touch special interests with something to hide.

The time will come when more powerful agents than the ORI's staff — congressional investigators, for example — will take a close look at standards of conduct in the life sciences in the United States. It is to be hoped that when this happens, scientific societies and the medical schools will be able to show that they have done more than reject scrutiny or accountability at every turn. ■

A turn-up for the National Science Foundation

A law passed by the US Congress shows unprecedented bipartisan support for basic scientific research at the NSF

This month's passage of legislation that allows for the doubling of support for the National Science Foundation (NSF) over the next five years does not guarantee that the money will actually be forthcoming. That will be determined in annual budget negotiations. But it does mark an unprecedented vote of confidence by the legislature in the NSF, and in the concept of supporting basic scientific research.

The NSF has always been a tough sell in the Congress. It doesn't distribute large sums of money to facilities around the country, and its research programmes — unlike those of, say, the National Institutes of Health — don't directly address real applications that are close to lawmakers' hearts. Rather, it distributes small grants to university researchers who propose the most scientifically interesting work.

It is therefore a remarkable accomplishment — and a great credit

to lawmakers who support the agency, such as House Science Committee chairman Sherwood Boehlert (Republican, New York) — that a law has been passed that may enable the NSF's annual budget to expand rapidly, from about \$4 billion now to \$8 billion by 2008.

The NSF earned this plaudit by consistently funding work on the basis of merit, and maintaining an honest peer-review system and an efficient management structure. Its director, Rita Colwell, has contributed to the momentum behind the bill by thinking big, and making a credible case that the agency can spend more money effectively.

If President Bush signs the bill into law, as expected, and then implements its provisions in his budget proposals, starting with the one released next February, he will truly have dispelled the notion that either his party or his administration is in any sense anti-science. ■





In the money
Congress backs
financial boost
for NSF
p257



Fair play
Revised misconduct
guidelines issued
for physicists
p258



Gene feast
Geneticists bowled
over by sequences of
rice chromosomes
p259



On guard
Conservationists
target wildlife
trade
p260

Biologists divided over proposal to create human–mouse embryos

Natalie DeWitt, New York

Human embryonic stem cells should be injected into mouse embryos to test the cells' likely clinical usefulness, some prominent US developmental biologists say. But the suggestion of creating such mixed-species embryos is likely to provoke public disquiet, and could galvanize political opposition to all research involving human embryos.

The idea was put forward on 13 November at a forum in New York, held to discuss standards for research with human embryonic stem cells. Ali Brivanlou, a developmental biologist at Rockefeller University in New York, and other supporters of the suggestion said that the 'chimaeric' embryos are needed to test the pluripotency of existing lines of human embryonic stem cells. Pluripotency — the ability to divide into cells of different functions, such as muscle or brain cells — is an important measure of the likely clinical usefulness of embryonic stem cells, a unique class of cells that can develop in different ways to perform many such functions.

But the forum, which was organized by the New York Academy of Sciences and Rockefeller University and included several US stem-cell researchers, failed to back a discussion document that would have included a call for the chimaeric embryos to be produced.

For years, mouse embryologists have tested pluripotency by injecting cells into a very early stage mouse embryo called a blastocyst, which consists of a hollow ball of cells. If the injected cells integrate into the embryo and contribute to the formation of every tissue, including the germ line which produces sperm and eggs, then they are pluripotent.

Ethical considerations would preclude an experiment that injected human cells into human embryos. But if they are compatible, cells from different mammalian embryos can be combined to form a viable chimaeric embryo. So some researchers are suggesting that the pluripotency of human embryonic stem cells could be tested by injecting them into mouse embryos. These embryos would then be reimplanted into a female mouse and allowed to develop. The tissues would be dissected and examined at various stages of development to see if they contained human cells.



All mixed up: should human embryonic stem cells be fused with mouse blastocysts (pictured)?

But the New York meeting ended without agreement on the issue. Ronald McKay, a leading stem-cell researcher at the National Institute of Neurological Disorders and Stroke (NINDS) in Bethesda, Maryland, said that he flatly opposed a draft document circulated at the meeting that contained the chimaera proposal. "I am completely opposed to putting human embryonic stem cells into any condition that will cause moral affront," he said.

James Battey, a developmental biologist at NINDS and chair of the influential National Institutes of Health (NIH) Stem Cell Task Force, was excluded from a closed session held at the meeting on the discussion paper, and criticized participants for what he regards as excessive secrecy. Battey added that the group, which included prominent researchers such as Fred Gage of the Salk Institute for Biological Studies in La Jolla, California, and Harvard's Douglas Melton, was not fully representative of the field. James Thomson of the University of Wisconsin at Madison, who first isolated human embryonic stem cells, and Austin Smith of the University of Edinburgh, UK, were invited but did not attend, and human embryonic stem-cell experts from Australia and Israel were not invited.

During the open part of the discussion, Janet Rossant of the Mount Sinai Hospital in Toronto questioned whether making interspecies chimaeras was scientifically informative, claiming that in such experiments, "success is meaningful but failure is not". For one thing, she said, the different gestation periods

for mice and humans may make it unlikely that the cells will combine in the embryo.

Rossant said there were viable alternatives to making interspecies chimaeras, such as assessing how the embryonic stem cells behave in culture, or testing whether they can engraft and form different tissues after injection into adult mice or mouse fetuses.

None of these tests would present the same ethical problems as producing chimaeras, Battey says. "Generally, the further along in development the fetus is, the less ethically complex these procedures are," he says.

Currently there are no NIH guidelines on any of these types of experiments. But an official at the agency says that they are generally viewed as ethically acceptable if they are done with embryos after the period of gonadal development has passed, so there was no danger, for example, of creating mouse sperm containing human DNA.

Experiments in which the cells of different species are combined in an embryo are likely to become more common as technology advances, researchers say. In September, Nissim Benvenisty of the Hebrew University of Jerusalem reported that he had grafted human embryonic stem cells into 1.5–2-day-old chick embryos (R. S. Goldstein, M. Drukker, B. E. Reubinoff and N. Benvenisty *Dev. Dynam.* **225**, 80–86; 2002). And Huizhen Sheng of Shanghai Second Medical University has transferred the genetic material of human cells into rabbit eggs (see *Nature* **419**, 334–336; 2002).

D. SPEARS LTD/SPL

Doubts linger over America's top herbicide

Rex Dalton, Salt Lake City

Rising concerns about the environmental impact of the most widely used herbicide in the United States have led researchers to call for more time to study its effects before approval for the agent's use is renewed.

Researchers have until next February to submit their findings on the herbicide atrazine to the Environmental Protection Agency (EPA). The agency plans to complete its review of the chemical during 2003.

But at the annual meeting of the Society of Environmental Toxicology and Chemistry (SETAC) in Salt Lake City, Utah, this week, some scientists said that the necessary studies cannot be finished in time for the EPA's deadline.

Researchers in the field want to conduct a detailed study of what concentrations of atrazine cause environmental damage by disrupting the sexual development of amphibians, says Timothy Gross, an ecotoxicologist with the US Geological Survey in Gainesville, Florida.

Earlier this year, two studies indicated that



Tyrone Hayes suggests that reasonably low levels of atrazine can adversely affect frogs' development.

atrazine at levels as low as 0.1 parts per billion (p.p.b.) in water disrupted the sexual development of male frogs (T. Hayes *et al. Proc. Natl. Acad. Sci. USA* **99**, 5476–5480; 2002 and T. Hayes *et al. Nature* **419**, 895–896; 2002).

But other scientists, including Gross, dis-

pute this, and have released unpublished results showing that much higher water concentrations of about 25 p.p.b. are needed to affect the frogs' sexual development. Government studies show that atrazine can be found at levels as high as 50 p.p.b. in US waters, and at even higher concentrations of several parts per million in agricultural run-off.

"The water is very muddy right now," says Gross, who also serves on a panel studying atrazine for Syngenta, the herbicide's Swiss-based manufacturer. Syngenta wants the EPA to re-register atrazine for use in the United States. It has been banned in some European countries.

Gross' comments came after a spirited session at the SETAC meeting where academic and industry-backed scientists presented their findings on the impact of atrazine on wild and laboratory amphibians.

Tyrone Hayes — an evolutionary endocrinologist at the University of California, Berkeley, who is the lead author of the studies in *Nature* and the *Proceedings of the National Academy of Sciences* — was added at the last minute to the SETAC session, which already included prominent university scientists, such as James Carr of Texas Tech University in Lubbock, whose studies of atrazine are funded by Syngenta.

Once February's deadline is reached, the EPA will set up a scientific advisory panel to advise it on re-registering atrazine. A decision is expected next October.

But both Hayes and Gross say that a new study is needed to settle the debate, and that the February deadline won't allow this to happen. "We need a tie-breaker study," says Gross.

Thomas Steeger, an EPA biologist working on atrazine re-registration, says that the agency may be reluctant to extend the study deadline. But he adds that he wouldn't be surprised if the scientific advisory panel called for more study before issuing a recommendation on how the herbicide should be regulated.

Cambridge-MIT Institute probed

David Adam, London

When the University of Cambridge, UK, and the Massachusetts Institute of Technology (MIT) announced plans for a collaborative institute (see *Nature* **402**, 111; 1999), the move was hailed as a major step forward for both universities. But three years on, the Cambridge-MIT Institute (CMI) is seeking new leadership, amid scant evidence of its results and recurring questions about why it was set up in the first place.

Both executive directors of the institute, which has been promised up to £68 million (US\$108 million) in British government funds, are to step down in January. A spokesperson for the CMI says that Alan Windle, the Cambridge executive director, and John Vander Sande, MIT executive director, are leaving to concentrate on their research, and that both feel they have finished their jobs of setting up the partnership.

The institute was billed as an attempt to bolster Cambridge's exploitation of its research by creating spin-off firms and building partnerships with companies. It was also intended to organize student exchanges and training programmes for academics, and to fund collaborative research projects.

But the project has been dogged by controversy. Critics have even dubbed it the 'Cambridge Dome' — a reference to London's Millennium Dome, which was built at great public cost to little apparent purpose.

"This was a vast sum of money given without proper supervision, and there was really no clear idea what to spend it on," claims Gillian Evans, a history researcher at Cambridge who was a member of the university's governing council and a critic of the arrangement.

The British government has refused to confirm what the institute's \$14-million annual funding allocation is being spent on, saying that it is not free to do so until the CMI, which operates as a limited company, publishes its accounts.

But in a report published on 6 November, the House of Commons Science and Technology Committee describes the decision to fund the institute as "somewhat curious". Ian Gibson, chair of the committee, says that he plans to question science minister David Sainsbury about it later this month. "It seems a strange way for money to come out of the blue on a whim of somebody," Gibson says. Other universities were angered by the decision to allocate the money to Cambridge without giving them the opportunity to compete for the partnership.

But defenders of the CMI say that with MIT wanting to collaborate with Cambridge, an open competition would have been a waste of time. Although the institute got off to a slow start, Vander Sande says that several research projects, joint courses and student exchanges are now taking place.

Congress backs historic expansion of NSF

Geoff Brumfiel, Washington

In a major boost for the physical sciences in the United States, Congress has passed a bill that recommends doubling the budget of the National Science Foundation (NSF) over the next five years.

Passage of the NSF Authorization Act of 2002, which President Bush is expected to sign before the end of the year, is seen as a historic landmark for the science agency, which supports most non-biomedical research at US universities, in disciplines ranging from computer science to geology.

Scientific societies have been eagerly pursuing such a measure for years, to match the doubling since 1998 of the budget of the main biomedical research agency, the National Institutes of Health (NIH). But it remains to be seen whether Congress will enact the 15% increases in the NSF's budget for each of the next five years recommended by the bill.

"Symbolically, this says there's bipartisan support for the NSF," says Samuel Rankin, chair of Coalition for National Science Funding, an alliance of scientific societies that lobbies for NSF funding. Congressional appropriations committees, which actually allocate funds, have already decided to give the NSF a 12% increase for next year, bolstering the view that they are ready to support the agency's rapid expansion.

If the doubling does take place, disciplines that are likely to benefit include Earth

sciences, mathematics, chemistry and computer science, which all rely heavily on the NSF for support. The agency also supports physics and biology, although the Department of Energy and the NIH, respectively, are far larger backers of these disciplines.

The doubling measure was welcomed by university physicists. "My hope is that this bill will allow the growth of new initiatives, while providing relief to ongoing programmes," says Barry Barish, a physicist at the California Institute of Technology in Pasadena and a newly elected member of the National Science Board, the NSF's governing body.

The bill was championed by Sherwood Boehlert (Republican, New York), chair of the House Science Committee. In a change from earlier versions of the legislation, the fourth and fifth years of increases are contingent on a congressional review of the NSF's management practices. This modification was one of several made in late negotiations with the White House Office of Management and Budget, which had previously opposed the bill on the grounds that five years was too long for a funding commitment.

The measure also directs the NSF to prioritize its large infrastructure projects. The agency currently has around half a dozen such projects awaiting funding, with no clear order in which they are to start. This situation is frustrating to project advocates in Congress, as well as researchers, says Rankin.



Sherwood Boehlert: champions funding hike.

Congress also passed a separate cybersecurity bill that recommends spending almost \$1 billion on computer security over five years, with half of the money going to the NSF. Gene Spafford, director of the Center for Education and Research in Information Assurance and Security at Purdue University in Indiana, says the cash would help research on the protection of advanced computer networks, and expand the number of doctorates issued in computer security. ■

ALEX WONG/GETTY NEWS

'Independent' biology institute targets China's exiles

David Cyranoski

China is to step up its efforts to lure talented biologists back home through a new institute that, its supporters say, will offer them more money and much more autonomy than other laboratories there.

After several false starts, the planned National Institute of Biological Sciences, Beijing, has just named its governing board, and over the next few months aims to recruit a director, 28 principal investigators and an external advisory board.

Bureaucrats often call the shots at China's main existing research institutes, creating a rigid image that can undermine efforts to entice back some of the thousands of Chinese scientists working abroad. "The new institute will be established as an independent organization and will be mainly run by biological scientists," pledges Ray Wu, a biologist at Cornell University in Ithaca, New York, and an informal adviser to the Chinese government on the project.

Researchers' salaries at the institute will be set at about halfway between typical

levels in the United States and in China — outbidding the prestigious Chinese Academy of Sciences (CAS), which has already had some success in attracting prominent scientists back to China. Each primary investigator will be allocated enough laboratory space for a team of ten researchers, postdoctoral fellows and technicians in the institute's building at the Zhongguancun Life Science Park, about five miles north of Peking University. The building, designed for a separate biological-sciences project, was taken over when those plans fell through.

The institute's investigators will mostly be recruited from abroad, says Wu, to avoid accusations of poaching talent from elsewhere in China. "The new investigators will add strength to the existing scientific community in China," Wu adds.

The project will receive 400 million yuan renminbi (US\$50 million) from the city of Beijing for building and equipment, plus 300 million yuan from the Ministry of Science and Technology to cover running

costs for five years. The institute will be devoted entirely to basic research, with its remit split between agricultural and health sciences.

"It will be able to recruit some of the best in the field," predicts Yongbiao Xue, associate director of the CAS's Institute of Genetics and Developmental Biology. "It could really push forward a lot of things in China."

Some researchers are sceptical about whether the institute will be able to maintain its promised level of autonomy, but many are optimistic. "The institute could set an example for research work in China," says Zhi-hong Xu, president of Peking University. "The system needs a lot of changes," he adds.

Xu and others are somewhat concerned, however, that insufficient provision has been made for interaction with other research institutions in China, and Xu suggests that the investigators should be required to hold joint positions with nearby universities. ■

Cancer researcher found guilty of negligence

Quirin Schiermeier, Munich

A university investigation has found the lead author of a contentious German cancer study guilty of gross negligence, but has cleared his 14 co-authors of misconduct.

The study, led by Alexander Kugler, formerly of the University of Göttingen, was published two years ago in *Nature Medicine* (A. Kugler *et al. Nature Med.* **6**, 332–336; 2000). It claimed that kidney cancer could be treated using a vaccine made from tumour cells fused with healthy dendritic cells from the immune system.

But the university launched an investigation into the paper after allegations that Kugler and another co-author had included a picture downloaded from the Internet in another paper submitted for publication (see *Nature* **412**, 8; 2001).

In a statement issued on 12 November, the University of Göttingen said that the Kugler paper “fails to meet the requirements of good scientific practice”. It added that data in the study were handled incorrectly, and that information relevant in judging the vaccine’s efficacy — such as whether the patients had received alternative treatments — was missing from the study.

Among other problems that were identified by the university’s task force was the false claim that the vaccinations had been carried out under a clinical trial approved by an ethical committee. In fact, they were individual treatments which were later subsumed into a study for publication.



But the task force firmly apportions the blame to Kugler, who has since left research, and whose whereabouts is unknown. The 14 co-authors, including the heads of the university’s institutes of urology and nephrology, Rolf-Hermann Ringert and Gerhard Müller, received only mild reprimands in the university’s statement. The task force’s full report has not been released publicly, pending comment from Kugler’s co-authors, but the university says that a copy is being sent to *Nature Medicine*.

“I am taking this very seriously,” says Beatrice Renault, editor of *Nature Medicine*, adding that she will need to read the report in full before deciding what course of action to take over the paper.

Despite doubts about the initial study, at least 400 patients have since been treated with the vaccine by Ringert and Müller. The results of these treatments have not been made public — but according to an unpublished analysis obtained by *Nature*, there were no significant tumour regressions among the 100 or so patients treated in Müller’s department. Ringert, who has treated more than 300 patients, declines to comment on the vaccinations, saying that the results have not yet been fully analysed.

Ulrike Beisiegel, vice-dean for research and ombudswoman at the University of Hamburg, criticizes the slow release of information about the case. “It is urgent that all relevant information be revealed,” she says. Beisiegel adds that the episode reflects deep-seated ethical problems in German clinical research.

Meanwhile, the DFG, Germany’s main research funding agency, has announced its own investigation into the Kugler paper. Reinhard Grunwald, the agency’s secretary general, says that it will seek to build on the university investigation and establish both the vaccine’s therapeutic effects and the degree of responsibility borne by all of the authors of the paper.

The Göttingen research was supported by Fresenius, a biotechnology and healthcare company that is based in Bad Homburg, Germany, which is expected to obtain approval shortly for a fresh study of the vaccine’s efficacy. ■

Physics guidelines drop equal-responsibility clause

Geoff Brumfiel, Washington

Responsibility for the integrity of a scientific paper need not always be carried by all the authors of the paper, according to ethical guidelines unveiled last week by the American Physical Society (APS).

The revised guidelines are the society’s response to recent physics misconduct cases, including that of Jan Hendrik Schön, a researcher at Bell Laboratories in Murray Hill, New Jersey, who was found to have falsified data in several high-profile papers and was sacked in September (see *Nature* **419**, 419–421; 2002). They contain an updated code of conduct and advice on how misconduct cases should be handled — but most significant is a revision of the duties of co-authors.

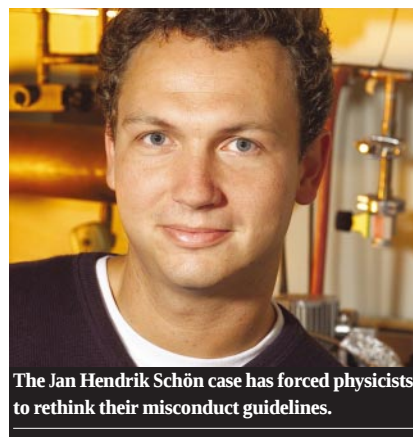
Previous APS guidelines stated that all authors held an equal share of responsibility for a paper. But several of Schön’s co-authors were cleared of misconduct because they had simply supplied him with materials.

The new code states that all researchers share “some degree of responsibility” for papers that they co-author, but only some have responsibility for the entire paper. “These include, for example, co-authors who are accountable for the integrity of critical data reported in the paper, carry out the analysis, write the manuscript, present major findings at conferences, or provide scientific leadership,” the guidelines say.

The move has the backing of many physicists, but also has critics, such as physicist David Goodstein, a vice-provost at the California Institute of Technology in Pasadena, who handles many of his institute’s misconduct cases. Goodstein says that fraudulent data almost always find their way into print before senior researchers are alerted, so it is unfair to hold supervisors accountable. “The senior people are responsible for bringing suspicions to the attention of the proper authority, but they’re not policemen,” he says.

William Brinkman, president of the APS, says that he disagrees with Goodstein but is open to modifying the guidelines in the future. “These are not like Newton’s laws — they’re not set in stone,” says Brinkman. ■

♦ www.aps.org/statements



The Jan Hendrik Schön case has forced physicists to rethink their misconduct guidelines.

BELL LABS

Fresh analyses put rice genomics on the map

Erika Check, Washington

Scientists who are trying to evaluate the usefulness of 'finished' genome sequences have fresh evidence to chew on this week, with the publication of high-quality sequences of two rice chromosomes.

The two papers, published in *Nature*, detail the complete sequences of rice chromosomes 1 and 4, with an accurate readout of which bases sit at each position in rice DNA (see pages 312 and 316 of this issue). The sequences also have very few gaps. Their authors say their analyses show how crucial 'finished' genome sequences are to academic researchers.

One group, led by Takuji Sasaki of Japan's Rice Genome Research Program, sequenced chromosome 1 and analysed it to find out how many genes it contains. The team calculates that the chromosome contains 6,756 predicted genes; a draft version of this portion of the rice genome released earlier this year predicted the existence of only 4,467 genes on chromosome 1. A group led by Bin Han of the National Center for Gene Research in Shanghai produced a finished sequence of chromosome 4, reporting that 52% of the genes were not completely predicted by the draft sequence.

But most importantly, scientists say, these completed sequences are linked to a physical map that shows where traits such as size and fertility occur on each rice chromosome. This is crucial for researchers who want to use genetic information to modify rice and other grains for agricultural uses. "It's incredibly important to have a map-based sequence for agriculture," says Ben Burr of Brookhaven National Laboratory in New York state, an adviser to the international, publicly funded Rice Genome Research Program.

Sasaki's group plans to have the entire rice genome finished to 'phase 2' completion within a month. Each piece of the rice genome will then have been sequenced an average of ten times and assigned a physical location.

The rice genome effort has been riddled with the same tensions over public data release as the human genome projects. Two companies, Monsanto and Syngenta, have produced 'draft' sequences of one rice strain, *Oryza sativa japonica*. Syngenta has not released all of its sequence data, but says that it has shared data with the international public project.

Stephen Goff, president of genome technology at Syngenta's Torrey Mesa Research Institute in La Jolla, California, says the company's draft is sufficient for their research. "Our interest is in gene discovery and getting good enough coverage to apply other genomics technologies, and the draft sequence works fine for that, so it didn't really serve our needs to go for the finished sequence," he says.

But researchers in the public consortium say that this leaves many academic scientists with a lot of work to do. "Every researcher has



Growing field: geneticists are excited by the prospect of highly detailed maps of rice chromosomes.

to go in and firm up the sequence," says Robin Buell of The Institute for Genomic Research in Rockville, Maryland, adding that although this may be easy for large industrial labs, it is much harder for individual university researchers.

This is important for rice, which many hope to use as a gold standard to aid their investigations of other cereals with larger genomes,

such as wheat, maize and barley. But it probably won't hold true for every genome that scientists want to sequence, researchers say.

"You'll find that some questions can be answered very well with a draft and some can't, and we don't know those trade-offs now," says Eric Green, scientific director of the US National Human Genome Research Institute. "That's what we need to learn." ■

Dinosaur curators plot mating game

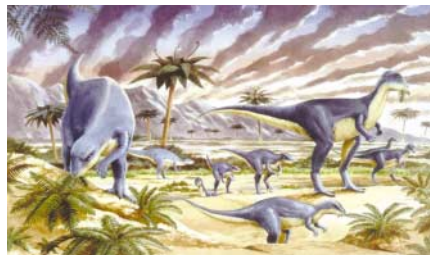
Rex Dalton, San Diego

Friends of 'Big Al', Montana's most famous dinosaur, want a buddy of the allosaurus to come home to the United States.

Big Al is displayed at the Museum of the Rockies in Bozeman. His buddy, a near-complete skeleton of another allosaurus, was stolen a decade ago from federal land in Utah and sold to a private museum in Japan.

Now Mark Goodwin — a palaeontologist at the University of California, Berkeley, who has helped to repatriate pilfered specimens in the past — has started a drive for the return of the rare dinosaur.

"The US government should seek the return of this skeleton," argues Goodwin. "If someone stole the Liberty Bell and it ended up in Japan, tell me it wouldn't be returned." He says that he plans to write to the US state department to argue his case.



Homeward bound? US researchers are calling for the repatriation of allosaurus fossil remains.

The purloined allosaurus is now at the Hayashibara Museum of Natural Sciences in Okayama, which is funded by Hayashibara, a large chemical firm. Museum officials have declined to offer to return the specimen, saying only that they are "ready to discuss the matter" with the US government. Shinobu Ishigaki, the museum's assistant director, adds that US palaeontologists are free to study the skeleton in Japan if they wish.

In July, Barry James, a Pennsylvania fossil dealer, pleaded guilty in a Salt Lake City courtroom to stealing the skeleton in 1992 and selling it to the Japanese museum for \$400,000. James was ordered to pay \$50,000 restitution and was placed on probation.

Federal authorities say that the Hayashibara Museum was duped into believing that the skeleton was found on private land. But to palaeontologists such as Jack Horner of the Museum of the Rockies, that shouldn't make a difference. Two years ago, for example, his museum returned a fossil skeleton to China, from where it had been improperly exported. "That is what a legitimate institution does," says Horner.

University of Wyoming palaeontologist Brent Breithaupt, who helped to rescue Big Al from the clutches of a commercial dealer in 1991, believes that diplomacy is the best way to recover the skeleton. "We need to see if we can develop a working relationship with the Japanese to get it back," he says. ■

Timetable for talks keeps hopes alive for bioweapons treaty

Geneva Hopes of strengthening the Biological and Toxin Weapons Convention were renewed last week, after signatories agreed on plans to tackle the convention's biggest stumbling blocks.

The future of the 1972 accord — which outlaws the manufacture, stockpiling or use of biological or toxin-based weapons — had been in doubt since November 2001, when the United States refused to discuss plans for laboratory inspections to assess compliance with the treaty (see *Nature* **414**, 675; 2001).

But at talks in Geneva last week, delegates put their differences behind them and agreed a timetable for discussing compliance issues and four other problems facing the treaty. These include the need to agree international rules for detecting and responding to alleged abuses of the convention, and the adoption of a code of conduct for scientists working with potential biowarfare agents. Delegates will meet annually to address each hurdle in turn. Another failure would have dealt a near-fatal blow to the treaty, postponing future discussion until the treaty's next review meeting in 2006.

John Walker, an arms-control expert at

Britain's Foreign and Commonwealth Office, describes the plan as a “decent work programme”, given that US negotiators refused to commit to firmer measures.

Wildlife protectors switch targets to mean business

Santiago The Convention on International Trade in Endangered Species (CITES) has come of age, assert conservationists who attended talks on the treaty on 3–15 November in Santiago, Chile.



In deep water: trade in seahorses must be controlled if they are to be kept from extinction.

CITES originally focused on the traffic in endangered plants and animals such as orchids and tigers, which is important to the species' survival but not hugely significant in economic terms. Now it has begun to tackle big wildlife businesses such as fishing and forestry. After a decade of arguing, the big-leafed mahogany (*Swietenia macrophylla*) was listed on CITES Appendix II, which covers species that are not at immediate risk of extinction but which still require trade to be closely controlled. All 32 species of seahorse were also placed on Appendix II. The animals are in demand for aquariums and traditional medicines.

Jim Armstrong, CITES' deputy secretary general, believes that the 160 nations signed up to the convention are showing a new willingness to apply it to “truly commercial” species. “It's quite a watershed,” he says.

Slump forces funding cut for Howard Hughes

Washington The Howard Hughes Medical Institute, the largest research philanthropy in the United States, is to cut its spending. The move is a response to the US economic downturn and stock-market slump, which has hit many charities (see *Nature* **419**, 765; 2002).

The institute's endowment has fallen from \$13 billion to \$10 billion in three years,

and is currently not producing enough income to support its core biomedical research programme, which accounts for 71% of its total annual spending.

The institute announced last week that it aims to reduce the budgets of its investigators by up to 10% a year for the next two years, review a grants programme for science education, and axe a \$22-million project to fund laboratory improvements at medical schools. It will also freeze recruitment and make some job cuts.

Russians brought to book as Newton snatch is derailed

Moscow A Russian couple have been arrested in connection with the theft earlier this month of a rare first edition of Isaac Newton's signature work *Philosophiae Naturalis Principia Mathematica*.

Police have declined to release the full names of the suspects, or the titles of the books reported to be in their possession when they were detained on 13 November on a train to the Caspian Sea port of Astrakhan. But the Interfax news agency quoted a police source as saying that a copy of *Principia* was found. Police say that the duo posed as graduate students to gain access to a reading room in the Russian National Library in St Petersburg, which later reported the book missing.

Newton formulated his laws of gravity and motion in the book, which was first published in London in 1687. Some 200 copies of the first edition still exist, each with an estimated value of around £200,000 (US\$316,000).

Royal Society fires up fossil-fuel tax debate

London The UK's Climate Change Levy, a government scheme aimed at cutting greenhouse-gas emissions, should be scrapped and replaced by a different tax or permit system, says the Royal Society.

The levy, which takes the form of a tax on non-domestic fuel and electricity, is not an effective way of cutting emissions, according to a report released on 18 November. Fossil-fuel use by households and transport is not covered, say the authors, and renewable electricity sources such as wind power, which do not produce greenhouse gases such as carbon dioxide, are unfairly penalized.

The society recommends replacing the levy with a tax on all fossil fuels, or a system of tradable permits that would be allocated to companies on the basis of cuts they make in their greenhouse-gas emissions. Such measures would make renewable energy sources more economically viable, the report suggests, and encourage energy conservation measures.

Terrorist's missing brain turns up in German lab

Munich The brain of left-wing terrorist Ulrike Meinhof, co-founder of the notorious Red Army Fraction who died in prison in 1976, was found at the University of Magdeburg in Germany earlier this month.

Meinhof and three other members of the Fraction, which carried out kidnappings and assassinations in Germany in the 1970s, died in disputed circumstances. Meinhof's brain was sent to neuropathologist Jürgen Peiffer at the University of Tübingen for forensic examination. Peiffer reported that the brain was damaged when a tumour was removed in 1962, which he argued could have accounted for Meinhof's aggressive behaviour. But few neuroscientists agree that behaviour can be

predicted from brain pathology.

Peiffer passed the brain to Bernhard Bogerts at Magdeburg University for further research in 1997. It has been kept there since then without the approval of the authorities or next of kin. The brains of the three other Fraction members have also vanished from the Tübingen clinic.



Ulrike Meinhof died in prison in 1976.

AP

The heavens at your fingertips

A small but growing group of astronomers wants to put the night sky on the Internet. But will staring at a computer screen ever replace peering through a telescope? Geoff Brumfiel logs on.

Pune in India is no place for an observatory. Located 160 kilometres south-east of Mumbai, formerly Bombay, the city is pounded by monsoon rains for four months of every year. Besides heavy cloud cover and high levels of starlight-distorting humidity, a telescope in Pune would have to contend with dust and light pollution from the city's three million or so inhabitants.

But Ajit Kembhavi, an astronomer at Pune's Inter-University Centre for Astronomy and Astrophysics, sees potential for a new kind of observatory there. "We can't hope to have large telescopes and instruments," he says. "But it would be great to have access to the data they produce."

Kembhavi is one of many astronomers with a similar dream. They want to construct 'virtual observatories' (VOs) — gateways to images obtained by the world's ground- and space-based telescopes. Electronic archives of astronomical images are already available. If VOs can overcome the technical and cultural difficulties of sharing that vast amount of data, they could help to democratize astronomy. "All of a sudden, anybody with an Internet connection has access to an incredible amount of knowledge and information," says George Djorgovski, an astrophysicist at the California Institute of Technology and a founder of the National Virtual Observatory in the United States. "It's very empowering technology."

The idea of an electronic star catalogue is nothing new. The Strasbourg Astronomical Data Centre (CDS), based at the Strasbourg Astronomical Observatory in France, was created in 1972 to catalogue the increasing amount of electronic data that are produced by astronomers. Other archives followed. NASA's Extragalactic Database began

archiving information on objects outside our Galaxy in 1987, and is now linked to data from the Hubble Space Telescope.

In recent years, the amount of online data has been further boosted by a new generation of sky surveys, which use digital cameras and computer programs to photograph, identify and file astronomical objects in huge online archives. The Sloan Digital Sky Survey, for example, is using a New Mexico-based telescope to chart the position of around a million galaxies by 2005 (see *Nature* **407**, 557; 2000). Other surveys have produced images of the entire Milky Way (see above), and new telescope projects could soon be making even more data available (see "Brace yourself for the data deluge", page 264).

Searching for progress

Researchers at the CDS began by cataloguing objects that appeared in the published literature. But as time went on, they and other archivers developed tools for analysing their databases. By using these techniques to compare images from the Sloan archive for example, researchers have revealed details about how the Universe is structured, such as the way in which galaxies cluster together. The archives and search tools are now immensely popular. Françoise Genova, director of the CDS, says that her database receives 10,000 hits a day — a huge number for a site that is used only by professional astronomers.

Peter Quinn, head of data management for the European Southern Observatory (ESO), says that the primary role of VOs will be to bring together the flood of data produced by the dozen or so sky surveys such as the Sloan, as well as those in the archives of satellites and ground-based telescopes. Three extensive VOs

are being developed — Europe's Astrophysical Virtual Observatory, headed by Quinn, the US National Virtual Observatory, and Britain's AstroGrid — with several smaller projects also taking shape (see table, opposite).

All are still currently at the prototype stage. Most are likely to use a single, easy-to-use interface to comb existing archives. In many cases, the researcher might not even be aware of all the databases that the VO is searching until it sends back the data. Some, such as the German Astrophysical Virtual Observatory, will also run their own databases, in this case of images supplied by the country's researchers.

Although VOs are still a long way from producing results, survey databases offer a taste of what they can offer researchers who lack access to cutting-edge equipment, says Stephen Landy, a physics professor at the College of William and Mary in Williamsburg, Virginia. "Researchers at William and Mary have no access to large telescopes," he says. Landy is interested in whether the Universe is flat — whether it obeys the normal rules of geometry. Studies of the large-scale distribution of galaxies should shed light on this issue, and Landy can do this analysis without a telescope by downloading the positions of thousands of galaxies from the databases of the Sloan survey and the Australia-based Two-Degree Field Galaxy Redshift Survey.

VOs, Landy says, will open even more doors for researchers at small institutions because they will provide a single access point to numerous databases. It will be "a hundred times easier" than using multiple uncoordinated databases, he predicts.

Combining different databases would also increase the possibilities for astronomical investigations. "The sheer size and quality of

these data sets lets you ask new kinds of questions," says Djorgovski. Searching for brown dwarfs, objects somewhere between a planet and a star in terms of size and mass, is one example. These dim objects emit optical and infrared radiation, so a simultaneous search of archives from optical and infrared telescopes would be a powerful tool for finding them.

In a similar way, trawls of X-ray, optical and radio databases will generate insights into galaxies with centres that produce huge amounts of radiation. These central regions emit radiation over a broader range of wavelengths than other objects, so many different types of telescope can be used to study them. Each wavelength has its own advantages and disadvantages. Radio waves, for example, are useful because they can pass through the dust that obscures some galactic cores, but they are emitted by only 10% of active galactic centres. Combining data from many wavelengths should give a much better picture of the properties of these galaxies, Djorgovski says.

Creating the tools needed to run such searches is one challenge faced by VO advocates. Software designed to pull information and correlations out of databases already

exists. "You can go on the web and buy software that does exactly this kind of thing now, except it's tuned to deal with hundreds or thousands of data points," Djorgovski points out. But to be effective, a VO must be able to sift through the spectral characteristics of billions of stars. "That is a major computer-engineering problem," he says.

Alex Szalay, an astronomer at Johns Hopkins University in Baltimore, Maryland, is working with other researchers to develop faster and smarter searching tools for the Sloan database, but admits that his collaboration has a way to go. "So far we haven't quite risen to the problem," he says. He hopes to improve matters by working with computer experts in public and commercial organizations, including Microsoft.

But developing search tools is not the biggest technical hurdle facing VOs. Because a patchwork of databases is involved, standardizing data quality is a significant challenge.

The resolution of an image depends upon the telescope used and the weather at the time it was taken. Likewise, different VO users will have different resolution requirements depending on their research. Accord-

New vision: Ajit Kembhavi believes that Pune is the perfect place to host a virtual observatory.

ing to Quinn, every image in the VO will have to include detailed information on the conditions under which it was obtained, so that researchers can decide whether the data are accurate enough to be used in their work.

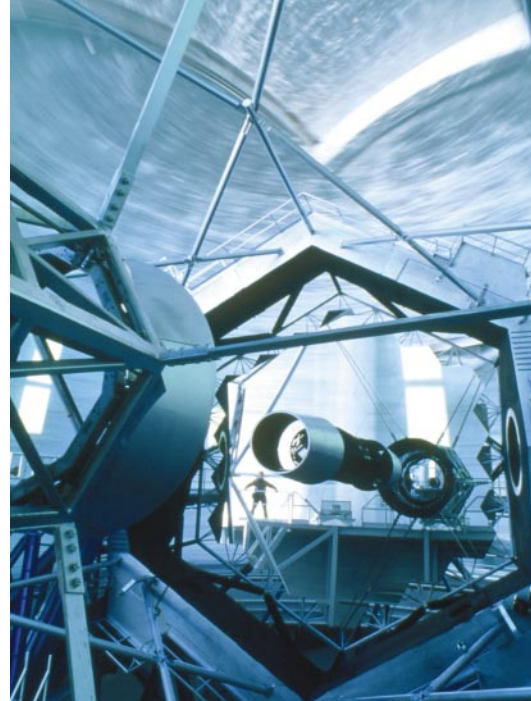
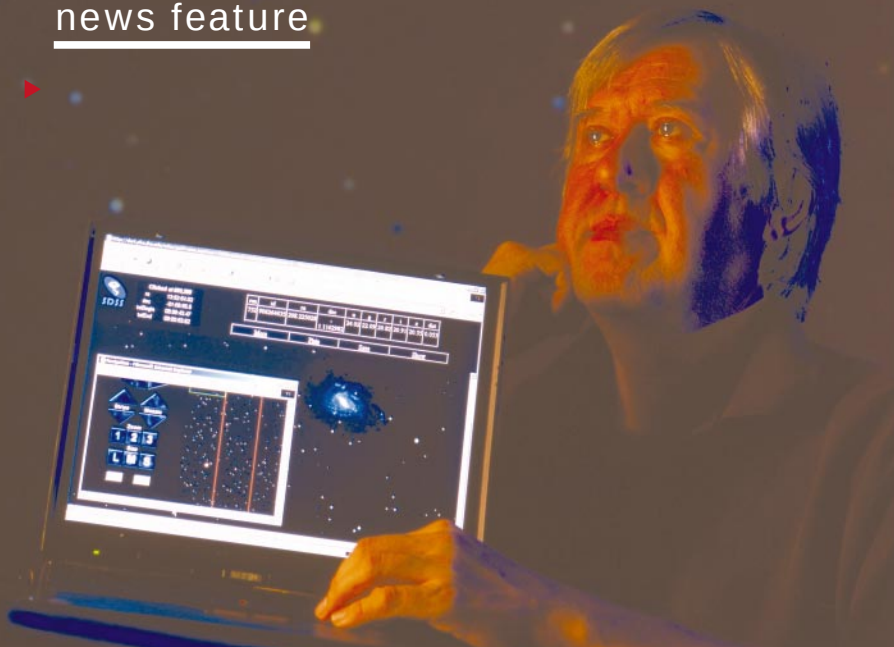
Data and time

Most sky surveys already attach such information to their images. But Sandra Faber, an astronomer at the University of California, Santa Cruz, says that it will be difficult to get astronomers into the habit of recording these 'metadata' when they add their data to databases maintained at telescope centres. Ground-based optical telescopes, for example, require lengthy calibrations, and recording the fine details of that process would be a drain on astronomers' often-limited observing time. "This is a very considerable overhead to the actual observers," she says. Some facilities, such as the ESO's Very Large Telescope at the Paranal Observatory in Atacama, Chile, already require users to record metadata. VO advocates are pinning their hopes on other observatories doing the same once VOs become more popular.

The problem of metadata underscores a much broader cultural challenge facing the VOs: astronomy has historically been a solitary activity. "In a typical ground-based observatory, if you get three nights on the telescope, those three nights are yours," says Andrew Lawrence, an astronomer at the University of Edinburgh and head of AstroGrid. So, adds Faber, are the data. "There is a general unwillingness to give away what you work hard to get," she says. Astronomers at

Starry line-up: the wave of virtual observatories being launched around the world

Project	Location	Funding (US\$ millions)	Grant duration (years)
National Virtual Observatory (VO)	United States	10	5
AstroGrid	United Kingdom	7.8	3
Astrophysical VO	Europe	3.9	3
Canadian VO	Canada	2.8	2
Indian VO	India	1.0	3
German Astrophysical VO	Germany	0.6	2.5
Australian VO	Australia	0.3	Not available
Japanese VO	Japan	0.3	1
Russian VO	Russia	0.01	3



Knowledge base: Alex Szalay hopes that facilities such as the Keck telescope (right), which currently has no archive, will develop searchable databases.

university-operated observatories are under no obligation to share their data with the wider community, for example. Those at public observatories are often required to share a copy of the raw data, but they are not necessarily required to provide the metadata that would make them useful to others.

In the United States, the situation is not helped by the fact that many of the largest and most modern observatories lack well-developed public databases. The Hawaii-based W. M. Keck Observatory has no archive, for example. And the publicly run facilities of the US National Radio Astronomy Observatory

store their data on digital data tapes that must be accessed manually upon request. Moves are being made towards creating online archives at both observatories. NASA, for example, plans to stipulate that all research at Keck sponsored by the agency should be deposited in a public archive. But progress is slow because of a lack of funding. "It really comes down to money," says Faber. "Money is needed to support the data-taking and the actual making of the archives, and then to make a usable user interface."

The cost of developing an archive should fall as they become more widely used, when the software can be copied or adapted from

existing databases. And as more centres sign up, facilities such as Keck may find themselves under pressure to set up archives. "Supposing that a whole bunch of these observatories really did get on board," says Faber. "What would happen to Keck? I think it would become a pariah."

But so far, the leaders of the VO projects say that the community's response has been lukewarm. "People tend to think, 'yes this looks good, tell me when it works'," says Lawrence. To try to maximize their popularity, VO leaders will initially aim to incorporate the most popular databases and tools into their software. If they succeed in capturing the attention of researchers, innovative research should follow. And if other researchers can see good results flowing from scientists who use VOs, they are likely to want to get involved themselves. "I think its going to be a community-enlightenment process to get this running," says Robert Hanisch of the Space Telescope Science Institute in Baltimore, who heads the US National Virtual Observatory.

But for the hundreds of astronomers at universities throughout India, the promise is too great to ignore, says Kembhavi. These researchers may never be able to gather funds for a world-class telescope, but a VO will mean that a telephone line is just as good. "Most of them can afford to buy a PC with their own money and for a very small sum they can access the Internet," says Kembhavi. "A virtual observatory will empower them to conduct relevant, high-quality research." ■

Geoff Brumfiel is *Nature's* Washington physical sciences correspondent.

US National Virtual Observatory

♦ www.us-vo.org

Astrophysical Virtual Observatory

♦ www.euro-vo.org

AstroGrid

♦ www.astrogrid.org

Brace yourself for the data deluge

The mass of data already produced by sky surveys could be dwarfed by the output of a proposed new telescope.

A consortium of US astronomers is currently developing plans for the 8.4-metre Large Scale Synoptic Survey Telescope (LSST). The \$120-million project will use a special three-mirror design to reduce distortion around the edge of its images, creating an exceptionally wide field of view. Together with an advanced digital camera, this will allow the telescope to catalogue 18 terabytes (1.8×10^{13} bytes) of data each evening, enough to survey the entire sky in only three nights, says project spokesman Tony Tyson, an astronomer at Bell Laboratories in Murray Hill, New Jersey.

The speed with which the LSST could scan the sky would give astronomers an almost real-time picture of the heavens, allowing them to spot unusual changes such as supernovae and near-Earth asteroids.

"There's a huge laundry list of exciting astrophysical science to do," says Tyson.

Detecting matter that does not emit light — 'dark matter' — is another item on the list. Dark matter is intrinsically difficult to study because it does not show up directly in an image of the sky. But dense clusters of dark matter exert a gravitational pull that can distort the light from galaxies behind them. By measuring these distortions, Tyson believes that the LSST can shed light on dark matter's

distribution throughout the Universe. In doing so, it could also provide a measure of 'dark energy', the mysterious force that seems to be pushing the Universe apart.

Within a decade, Tyson says, the LSST could catalogue 15 petabytes (1.5×10^{16} bytes) of data, enough to fill about 1.5 million compact discs. That will be far too much for LSST collaborators to hoard, he says. So the project calls for the group to put all of their data online almost as quickly as they can record them — a promise of openness that VO advocates hope will soon be common.

Tyson says that the LSST team plans to raise funds from private and public sources, and hopes to complete the project within ten years.

Immunity's pregnant pause

If we can understand why a woman's body does not reject her fetus, it could help us to treat infertility and prevent problems in pregnancy. Helen Pearson reports.

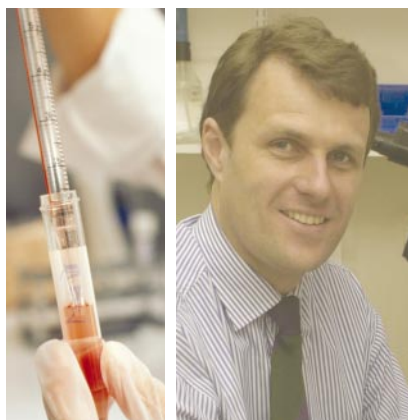
BSIP LAURENT/AUBESPL

In 1989, a young British woman had her ninth consecutive miscarriage. Her marriage broke down shortly afterwards. But within months of finding a new partner, she had conceived again and the pregnancy went without a hitch. Her daughter is now a healthy and lively nine-year-old.

Reproductive immunologists suspect that the woman's immune system took offence at her first choice of partner — over-reacting to tissues carrying his genes and expelling the fetuses he fathered. According to some experts, infertility, recurrent miscarriage, premature delivery and a dangerous complication of pregnancy called pre-eclampsia may all, in some cases, be linked to immunological abnormalities.

If these scientists are right, immunology could be reproductive medicine's next frontier, helping to treat distressing conditions that blight the lives of many couples. "This will be the new area of infertility treatment for this century," predicts Kelton Tremellen of the University of Adelaide in Australia.

In fact, how any fetus survives gestation has baffled scientists for decades: the embryo's tissues are half foreign and yet, unlike a mismatched organ transplant, it isn't normally rejected. Now researchers are revealing how different arms of the mother's immune system react to semen and the implanting embryo — and how the placenta protects itself from her immune system's attack. Some are already building from these findings to develop treatments that could help difficulties in conception and pregnancy. But there is also concern that rogue clinics may exploit prelim-



Kelton Tremellen says IVF (left) may be improved by preparing the way with proteins from semen.

inary results to offer couples therapies that have yet to be proved safe and effective.

The idea that the immune system may be to blame for problems with pregnancy stems in part from a long-standing observation about pre-eclampsia. This is a life-threatening condition for mother and child affecting up to one in ten pregnancies, in which the placental blood supply is insufficient and can starve a fetus of oxygen. First pregnancies are more susceptible to pre-eclampsia — unless a woman switches partner, in which case a second pregnancy is at equal risk. Previous exposure to a fetus carrying a particular suite of paternal genes, it seems, makes the immune system more likely to tolerate the first-born's subsequent siblings.

More recently, immunologists have wondered whether exposure to proteins in semen

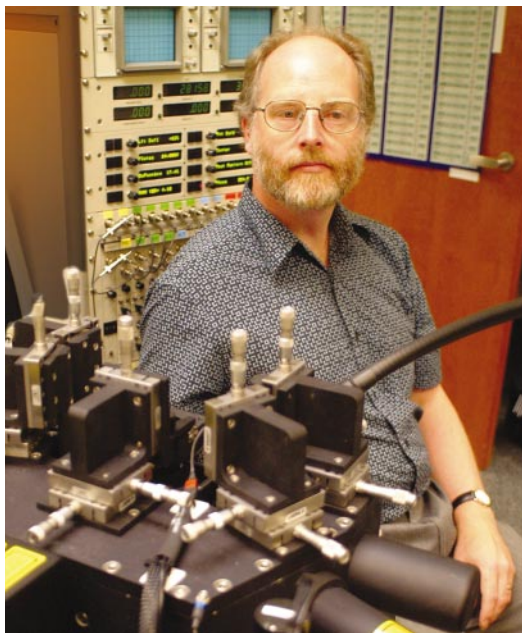
helps to prepare a woman's immune system for conception and pregnancy. Tremellen and his colleagues have studied one such protein, called TGF- β , found at high levels in semen. They injected TGF- β into a mouse's uterus alongside a cocktail of foreign proteins, and found that later injections of the same proteins under the skin did not elicit a strong immune reaction¹.

Tremellen, a gynaecologist working in the lab of reproductive immunologist Sarah Robertson at Adelaide, believes that 'immunization' with TGF- β through sexual intercourse helps the maternal immune system learn to tolerate molecular signatures, or antigens, in semen by altering the production of inflammatory molecules called cytokines. He has already shown that *in vitro* fertilization (IVF) is more successful if couples have sex beforehand², and hopes soon to begin trials of a vaginal TGF- β gel to see if it can help women suffering recurrent miscarriage.

Although proteins in semen may smooth the way for a future baby, damping down a mother's immune response to the implanting embryo also seems to be crucial to a pregnancy's success. At implantation, a group of embryonic cells called the trophoblast aggressively invades the mother's uterus lining, anchoring the placenta and widening the maternal arteries to enhance its blood supply.

Ashley Moffett-King of the University of Cambridge, UK, and her team have found that natural killer (NK) cells flood the uterus at this time. These immune cells are usually involved in attacking cancerous or virus-infected cells. But those drawn to the uterus

Z. MACAULAY/GETTY IMAGES



Andrew Mellor (left) and George Chrousos are tracing how a fetus hampers its mother's immune system.

at implantation carry receptors that interact with two unusual antigens, HLA-C and HLA-E, on the surface of trophoblast cells.

Exactly what happens as a result of the interaction between NK cells and the trophoblast is still subject to debate³. Moffett-King suspects that it triggers the production of particular cytokines that either help the trophoblast to invade the uterine wall or limit the extent to which it invades. When this process goes wrong, she suggests, the fetal blood supply is compromised and pre-eclampsia might result. If so, a mismatch between HLA antigens on the fetal cells and receptors on the mother's NK cells might predispose certain women to the condition. Moffett-King is now embarking on a study to search for susceptible combinations in 300 British women.

Fetal distraction

Other researchers have discovered several ways in which embryonic tissues disable the mother's immune system — including secreting suppressive factors into her blood or displaying them on placental cells. Another antigen called HLA-G, for example, is released from the trophoblast into the mother's bloodstream, and seems to protect the trophoblast from attack. Researchers led by Philippe Le Bouteiller, who works for INSERM, France's medical research agency, at the Purpan Hospital in Toulouse, have recently shown that soluble HLA-G makes certain types of immune-system T cell — which may attack fetal cells bearing the father's antigens — commit suicide⁴. What's more, a team led by Olavio Baricordi at the University of Ferrara in Italy recently found that implantation after IVF only occurred if the embryos secreted soluble HLA-G (ref. 5).

Based on such findings, Joan Hunt of the University of Kansas Medical Center in Kansas City has already patented a genetically engineered form of human HLA-G. "It could be administered to women with fertility problems," she suggests.

Last year, George Chrousos of the National Institute of Child Health and Human Development in Bethesda, Maryland, presented further evidence that a developing fetus induces some of its mother's T cells to commit suicide. His team showed that corticotropin-releasing hormone (CRH), which is secreted by both the implanting embryo and the lining of the uterus, stimulates trophoblast cells to produce a molecule, known as the Fas ligand, that binds to a cell-surface receptor that triggers cell death. When T cells were grown together with trophoblast cells stimulated by CRH, they died⁶.

Later in pregnancy, CRH seems to have an inflammatory role that helps to trigger labour. Chrousos has identified a molecule called antalarmin that blocks the action of CRH (ref. 6), and in unpublished work he found that it seems to delay labour in sheep.

Andrew Mellor of the Medical College of Georgia in Augusta, meanwhile, believes that the placenta cripples maternal T cells in another way: by starving them. In 1998, his team showed that the mouse placenta makes an enzyme called indoleamine 2,3-dioxygenase (IDO), which breaks down an amino acid called tryptophan that is vital for T cells' nutrition. Injections of a molecule that blocks IDO caused high rates of abortion in mice, the team found⁷. "The simple prediction is that problem pregnancies are associated with variations in the IDO gene," says Mellor — although its role in human pregnancy has yet to be demonstrated.

More recently, Mellor's team has shown that the fetal rejection triggered in mice by blocking IDO activates an arm of the immune system called complement⁸. This series of blood plasma proteins is one of the first lines of defence against bacterial and viral infections. But Hector Molina of Washington University in St Louis has shown that it can also hinder pregnancy. Molina's group found that mice lacking Crry, a cell-surface protein that protects tissues against misdirected complement activation, lose their fetuses⁹. Molina now plans to study whether women suffering recurrent miscarriage or pre-term labour have altered levels of proteins that mediate complement, or mutations in the human equivalent of the gene for Crry.

Birth control

With each researcher advocating their favoured mechanism for sustaining gestation, it will be some time before a complete picture emerges. But most agree that mother and fetus use several means in parallel to bring about a peaceful nine months. "There's no definite gene required for gestation," says Mellor. "But if you take one out there's more risk that a pregnancy will be lost."

The eventual goal is to convert the growing knowledge of reproductive immunology into treatments for infertility, miscarriage and other complications of pregnancy. Here, the field's pioneers are battling not only scientific and clinical complexities, but also practitioners operating at the fringes of reproductive medicine who are willing to offer highly speculative 'treatments' to couples desperate to have a baby.

Since the 1980s, for example, women suffering recurrent miscarriage have been offered injections of their partner's white blood cells, supposedly to prime their immune system into accepting a fetus bearing his antigens. Researchers have protested that this is not based on good evidence, and they have been vindicated by a negative randomized clinical trial¹⁰.

Disturbingly, some US doctors are now offering diagnostic tests to determine whether the NK cells in the mother's circulation are abnormal, even though this may be of little relevance to the interactions in the placenta being studied by Moffett-King.

"People try to get the clinical solutions before the science is worked out," says Ian Sargent, an immunologist at the University of Oxford, UK. "It has dogged this area."

Helen Pearson works in Nature's news syndication team.

- Robertson, S. A., Ingman, W. V., O'Leary, S., Sharkey, D. J. & Tremellen, K. P. *J. Reprod. Immunol.* **57**, 109–128 (2002).
- Tremellen, K. P. *et al. Hum. Reprod.* **15**, 2653–2658 (2000).
- Moffett-King, A. *Nature Rev. Immunol.* **2**, 656–663 (2002).
- Fournel, S. *et al. J. Immunol.* **164**, 6100–6104 (2000).
- Fuzzi, B. *et al. Eur. J. Immunol.* **32**, 311–315 (2002).
- Makrigiannakis, A. *et al. Nature Immunol.* **2**, 1018–1024 (2001).
- Munn, D. H. *et al. Science* **281**, 1191–1193 (1998).
- Mellor, A. L. *et al. Nature Immunol.* **2**, 64–68 (2001).
- Xu, C. *et al. Science* **287**, 498–501 (2000).
- Ober, C. *et al. Lancet* **354**, 365–369 (1999).

Footprint: ignoring the facts that don't fit the theory

The eco-sceptic idea that wealth leads to sustainability overlooks all reasonable evidence.

Sir — Your News report “Ecological footprint forecasts face sceptical challenge” (*Nature* **419**, 656; 2002), stating that Bjorn Lomborg’s environmental-policy group disputes the usefulness of the ‘ecological footprint’ concept, comes as no surprise to those familiar with Lomborg’s recent work. He argued in *The Skeptical Environmentalist: Measuring the Real State of the World* (Cambridge University Press, New York, 2001) that societies become “greener” and presumably more sustainable as they accumulate wealth and purchasing power. To support this ‘wealth hypothesis’ he correlates his preferred measure of environmental impact with gross domestic product (GDP) by country. However, his preferred measure is the Environmental Sustainability Index (ESI) generated by the World Economic Forum (WEF).

This index, plagued with methodological problems, offers no useful information about our long-term affairs and prospects. It considers 67 variables divided among 22 environmental indicators, most of which have little to do with a region’s ability to provide ecosystem goods and services.

For example, high petrol prices and the memberships of environmental organizations are tabulated, as are the number of scientific articles a country publishes and its record of regulatory innovation — but these probably do not relate to the general workings of the biosphere.

Additionally, none of the variables is weighted appropriately, so ecologically tenable variables such as soil degradation, water availability, ecological deficit and fertility rate are considered equal to petrol prices and measures of political corruption. To its credit, the WEF report recognizes these limitations, some of which may be addressed in future versions of the index.

The ecological-footprint concept offers a snapshot of our resource demand independent of GDP — unlike the ESI. Thus, by recognizing the reality of biophysical limits, a country’s ecological footprint offers a conservative measure of the total productive land, water and resources required to support a given population and cope with its waste materials. Such measures are necessary before we can hope to deal with more

difficult concepts such as regional and global sustainability. Lomborg’s group misses the point by stating “there is no ‘correct’ ecological footprint”. No one has claimed that an ecological footprint calculation is “correct” and should be accepted uncritically.

Estimating anything will always be subject to error — a fact that a statistician such as Lomborg surely recognizes. But given the urgent need to understand fully our global predicament and true prospects for sustainability, should we abandon conceptual tools if they rely on estimates?

I am aware of no measure of ecological footprint that supports Lomborg’s wealth hypothesis, which may explain why he dismisses the value of the ecological-footprint concept. It is not clear to me how Lomborg’s view will contribute to sustaining a growing human population over the long term, given the challenges of becoming regionally sustainable.

Keith M. Vogelsang

Department of Biology, Jordan Hall 142,
Indiana University, Bloomington,
Indiana 47405-3700, USA

Footprint: our impact on Earth is getting heavier

Sir — Bjorn Lomborg’s comments as reported in your recent News story (*Nature* **419**, 656; 2002) betray ignorance of ecological-footprint analysis (EFA). This is a method I proposed (W. E. Rees *Popul. Environ.* **24**, 15–46; 2002) specifically to assess the assertions made by some economists that, because of technological advances, the human economy is “dematerializing” or “decoupling” from the natural world. If true, this implies that modern society is becoming less dependent on nature, and can be used to justify further economic growth. If false, the additional stress likely to be imposed on the natural world will be disastrous.

Certainly, no single index can represent the total human impact on the ecosystem. However, EFA is comprehensive enough to show, unambiguously, that the human eco-footprint on Earth is steadily increasing. Comparative studies (such as the Worldwide Fund for Nature’s *Living Planet Report 2002*, criticized by Lomborg) show that the most technologically advanced nations are the most energy- and material-intensive and have the largest per-capita ecological footprints. Hence,

the consumer lifestyles of their average citizens (and of the wealthy residents of the developing world) are the least ecologically sustainable on Earth, and cannot be safely extended to humans everywhere.

Lomborg’s ignorance of eco-footprint science is illustrated by his claim that it ignores the potential growth in the use of renewable resources, and that if the use of such resources increases, the 2050 footprint would not be as large as projected. The fact is that EFA is already largely based on use of renewable resources, with the exception of the fossil-fuel (carbon sink) component. A shift to exclusive use of renewable energy does not necessarily imply a reduction in the eco-footprint.

Biomass fuels, for example, are likely on thermodynamic grounds alone to demand a larger growing area than the amount of land needed to provide the energetically equivalent amount of fossil energy (via carbon assimilation) today. One proposed ‘renewable’ resource, fuel ethanol production, is not even a net source of energy. In the United States, 70% more energy (mostly fossil fuel) is consumed in producing a litre of ethanol than is contained in the product. Even without counting the energy cost of the fermentation and distillation process, or the carbon-sink footprint of the fossil fuel used to grow the maize feedstock, the

average US automobile would require 4.4 hectares of cropland to provide its annual fuel requirements. (This whopping — but only partial — fuel eco-footprint is about seven times the cropland needed annually to feed one person in the United States.) To run all US cars on ethanol would require an area of cropland equivalent to nearly the total US land area. More generally, the United States uses 85% more fossil energy each year than the total energy captured by all its plant biomass over the same period. Clearly, regardless of the efficiency of the conversion technology, there is no possibility of renewable biomass substituting, joule for joule, for fossil fuel.

The much-touted solar alternatives may not fare much better. Although the solar flux represents a vast flow of potential energy, various analysts argue, from both thermodynamic principles and empirical data, that humanity faces enormous technological obstacles in converting this flow into the energetic equivalent of contemporary fossil-fuel use.

In the light of such data and the implicit uncertainty about prospects for sustainability, Lomborg’s “sceptical challenge” to the ecological footprint concept falls flat. EFA may be static, but this does not invalidate estimates of future footprints based on reasonable

assumptions about technological development. Dynamic modelling is just as vulnerable to implicit error in this regard.

William E. Rees

*School of Community and Regional Planning,
University of British Columbia,
6333 Memorial Road, Vancouver,
British Columbia V6T 1Z2, Canada*

Xenotransplantation's benefits outweigh risks

Sir — Your News item “Diabetes trial stirs debate on safety of xenotransplants” (*Nature* **419**, 5; 2002), discussing our clinical trial of pancreatic islet xenotransplantation, states: “Mexico has no published guidelines on clinical trials in xenotransplantation.” This is untrue. Mexico has a general health law that regulates the conduct of all experimental investigation in human beings; it also regulates organ and tissue transplantation, and xenotransplantation, which was added recently. The protocol for our clinical trial was approved by the research, ethics and biosafety committees of our institution and of the National Autonomous University of Mexico Medical School.

Your article did not mention the existence of other published studies on porcine–human islet xenotransplantation, although these studies are admittedly small, use different methods from ours and do not show significant benefit to patients. With regard to your reporter’s comment about the need to test “any proposed approach” on non-human primates, there are many examples in the fields of both transplantation and diabetes where experimental studies in humans are performed following evidence of benefit derived only from other mammalian models. Indeed, one of the problems we faced was that there is no appropriate model for auto-immune diabetes in non-human primates.

You report our critics’ views that we should have undertaken preclinical studies before proceeding, but vast amounts of preclinical data on animal models already exist, both on efficacy of islet xenotransplantation and the protective effects of Sertoli cells, including our own published work. We have not published some of our work, to protect patents. But we consider the available literature sufficient to justify our trials on young diabetic patients, the population that stands to benefit most.

You report our critics as stating that our work was on patients too young to give proper informed consent. We recruited adolescent and high-school students because they had had diabetes type 1 for more than four years, so that spontaneous remission would not occur, yet there were no diabetic complications typical of adult

patients. It is our strong contention that they were old enough to understand fully the potential risks and responsibilities. We had several meetings with the patients and their parents to explain in detail the procedure. After time for reflection, they signed extensive informed consent forms conforming to international requirements, for example the Helsinki declaration.

Type 1 diabetes causes the death of more than 40% of patients before they reach the age of 40, and is a leading cause of terminal renal insufficiency and blindness, among other devastating complications. We believe that the benefits of better metabolic control, and insulin and immunosuppression independence, offered by pancreatic islet xenotransplantation far outweigh the risks.

Rafael Valdes Gonzalez

*Hospital Infantil de México “Federico Gomez”,
Laboratorio de Xenotrasplantes, Edificio Mundet
4° Piso, Calle Dr Marquez 162, Col. Doctores,
CP 06720, México D.F.*

UK government closes its eyes to medical needs

Sir — Readers of *Nature* may get the impression that funding for basic research in the United Kingdom is in good health (“Calling for entrepreneurs: London” *Naturejobs* 19 September, 4–5; 2002). We wish to put the record straight.

The UK government uses the periodic Research Assessment Exercise (RAE) to determine levels of funding for university departments and institutes. Following the latest RAE, funding for hospital-based clinical subjects, which underpins the core of translational research in medical schools across the United Kingdom, was severely cut. Despite having received, for the third consecutive time, the highest possible rating (5*), the Institute of Ophthalmology in London faces a 19.7% cut in its predicted annual budget because the government has failed to honour its pre-RAE fiscal commitments. The problem is even greater for those who received a rating of 5 and who now face a massive 38% cut.

This takes £970,000 (US\$1.5 million) a year out of our operating budget. We are unable to appoint even a single technical post that we require to run newly commissioned laboratories, which puts extra stress on academic staff. The cuts also exacerbate the financial difficulties faced by our parent body, University College London, after years of underfunding.

The reason UK medical schools are having problems is self-evident, and the crisis extends across the country. In the longer term our institute, together with others in hospital-based clinical subjects,

will be less able to fulfil its mission, which is to deliver the translational research that enables laboratory science to be developed into clinical trials and ultimately into benefits to patients.

The message appears to be that, with regard to translational medical research, the government is not only blind to the achievements of 5* departments and institutes, it has also, incomprehensibly, decided to slash the support they receive. We also suspect that the cuts are due at least in part to inept planning and muddled thinking. Institutes such as ours have to budget, recruit and develop a coherent research strategy within the constraints of government policy. It is therefore essential that the government keeps its promises. This it has not done. In the United Kingdom the phrase ‘science policy’ has become an oxymoron.

If, as I. B. Holland suggests (*Nature* **419**, 248; 2002), underfunding is less acute here than in other European countries, then the prospects for our continent of beleaguered researchers must be bleak. Governments are often short-sighted when it comes to policy development, but in this case we are confronted by a total lack of vision. And as we are only too well aware at this institute, myopia may be corrected but blindness remains virtually untreatable.

**Stephen E. Moss, Gary Rubin,
John Greenwood, Adam Sililo**

*Institute of Ophthalmology, University College
London, 11–43 Bath Street, London EC1V 9EL, UK*

Mother knows best

Sir — In his otherwise excellent News and Views “The grand assault” (*Nature* **419**, 493–494; 2002), Russell F. Doolittle writes: “Eukaryotes can be loosely defined as organisms whose cells have nuclei and cytoskeletons, distinguishing them from the Bacteria and the Archaea, neither of which has introns in their coding sequences.”

Bacteria and the phage that infect them do contain introns in their genomes. The first example of an intron in a bacterial system was found in the thymidylate synthase gene of bacteriophage T4 (F. K. Chu, G. F. Maley, F. Maley and M. Belfort *Proc. Natl Acad. Sci. USA* **81**, 3049–3053; 1984) — the last author being my mother. Since then, hundreds of introns have been found in archaea, bacteria and their phage.

It has been a long time since I uttered these words, but I couldn’t be more proud to say: “Mommy told me so.”

Gabriel M. Belfort

*Department of Biochemistry, K-building,
Boston University School of Medicine, 715 Albany
Street, Boston, Massachusetts 02118, USA*

Fetal positions

The ethical opposition to cloning technologies appears to be softening.

Making Babies: Is There a Right to Have Children?

by Mary Warnock

Oxford University Press: 2002. 126 pp.
£9.99, \$13.95

Cloning

coordinated by Anne McLaren

Council of Europe Publishing: 2002. 190 pp.
15 euros, \$23

Dorothy C. Wertz

Attitudes towards human reproductive cloning may be quietly shifting among bioethicists. Mary Warnock, who chaired the committee on human fertilization whose report led to Britain's Human Fertilisation and Embryology Act (1990), once supported an outright ban. But she now believes that cloning, if ever proven safe, might be morally acceptable for complete male infertility that cannot be treated in any other way.

Using simple, straightforward, logical arguments aimed at the average person, Warnock demonstrates that other means of family formation that stand outside the usual norms are also morally allowable, because they have not been shown to cause greater harm than ordinary means of reproduction and may answer deeply felt needs of prospective parents. She begins with assisted reproduction for homosexuals, arguing that ethics should not be based on the fear of transgressing an imaginary, partly nineteenth-century Romantic line between 'natural' and 'unnatural'.

Absolute bans on any form of reproduction do little good, because they may have to be revised in the light of social experience. Warnock now believes that the 1990 UK ban on commercial surrogacy was a mistake, because today's private, 'unpaid' arrangements cannot be regulated or overseen, nor can the parties be counselled. It would be better to leave unusual reproductive arrangements in the open, regulated by a non-political government agency with some flexibility, along the lines of Britain's Human Embryology and Fertilisation Authority, rather than to react with absolute bans.

Warnock draws the line (as do many ethicists) at a moral and legal right to public funding for assisted reproduction, but argues that the duty of compassion in medicine means that people should be served if public resources allow or if they can pay out-of-pocket expenses. She fears that the concept of rights will be transformed into a popular belief that "I must get my due", which could undermine the foundations of public health insurance.



Be my baby: it may be better to regulate unusual reproductive arrangements than to ban them.

Although the book as a whole is clearly structured and logically presented, Warnock's arguments about reproductive cloning exhibit some confusion. Why should it be morally acceptable for male infertility, but not for female infertility or the replacement of a dead child or a miscarried fetus? Having said that ethics should not be based on fear, Warnock bases her opposition on the fear that dictators will clone armies of supporters, a ludicrous proposition in view of the numbers of women required to bear these children and the years needed to raise them. Nevertheless, the book is sprightly reading and food for serious thought. It would provide a good basis for discussion in high schools, colleges and book clubs.

Anne McLaren's edited volume presents a broad range of scientific, legal, theological and philosophical views on cloning from Britain, France, Germany and the Netherlands. These are presented in considerable depth and cover both reproductive and 'therapeutic' cloning, in which stem cells are made to order from the patient's own adult cells. The book also contains appendices explaining the technique of somatic-cell nuclear transfer, a list of websites, and the text of the January 1998 Council of Europe Additional Protocol to the Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine, on

the Prohibition of Cloning Human Beings, which banned the cloning of human beings. The Council of Europe left it to individual member states to define 'human being'. Most of the larger European nations have not yet ratified the European Union accords.

Many of the scientific terms have been thoughtfully explained for non-scientists in marginal notes. However, most of the chapters are written at a level of ethical scholarship that is more likely to appeal to professionals than to the public, although some chapters could be used in universities.

The authors' views are varied. Some take a hard line against all types of cloning, on the basis that the blastocyst, a ball of about 100 cells that appears about 6 days after conception, is a 'human being'. Others think that reproductive cloning is not an affront to human dignity and should be allowed if proven safe. Still others think that both reproductive and therapeutic cloning are impractical and will never occur.

John Gurdon and Keith Campbell, two pioneers of animal cloning, describe the history and science of cloning and strongly support therapeutic cloning, as does the biologist Axel Kahn. But the theologian Dietmar Mieth opposes therapeutic cloning, on the traditional Catholic grounds that the embryo is a human from the moment of conception and on the kantian grounds that the use of a human as a means

GANDEE VASAN/GETTY

to even a beneficial end is an affront to human dignity. McLaren, a biologist and member of Britain's Human Embryology and Fertilisation Authority, thinks that therapeutic cloning is impractical, unlikely and unrealistic because it requires too many human eggs.

Protestant theologian Egbert Schrotten attacks the problem of human dignity head-on and concludes that human reproductive cloning is not intrinsically wrong. The moral status of a human embryo depends on its intended use (to be discarded or placed in a womb), rather than on its ontological status. The 'unnaturalness' of cloning is not an argument against its use, he argues: all technology is artificial and thus by definition unnatural. Nature is not intrinsically good and does not teach us how to behave. Schrotten sees no reason to prohibit research, but would place it under the supervision of a politically independent, multidisciplinary licensing body.

The obstetrician Claude Sureau, former president of the French National Academy of Medicine, after presenting in detail several scenarios from science fiction and clinical reality, including a dying spouse, male (but not female) infertility, and mitochondrial disease, concludes that "reproductive cloning is more a figment of the popular imagination than a genuine threat to humanity". He believes that psychosocial outcomes are more important than "any ideology or dogma-ridden theoretical speculation about the nature and status of the embryo". Writer Colin Tudge urges humility because we do not know what the effects (or non-effects) of cloning may be.

Maxime Tardu provides a legal review, but seems to think that the US Senate has already passed a law against all cloning, including therapeutic cloning. However, this has not yet happened and is unlikely in this session of Congress. Tardu calls for a moratorium for further reflection rather than an outright ban. Andre Albert of the French Ministry of Justice thinks that a worldwide ban on reproductive cloning, probably through the United Nations, is necessary to preserve human dignity. But he does not discuss how such a ban might be enforced.

If this book had been written in 1997, when the birth of Dolly the cloned sheep was announced, most chapters would probably have opposed reproductive cloning. Now that we have had time to reflect, reproductive cloning appears less abhorrent, and even acceptable to some. Historically, this shift has been the first step on the road to acceptance for other assisted-reproduction technologies. ■

Dorothy C. Wertz is at the University of Massachusetts Medical School, Shriver Division, 200 Trapelo Road, Waltham, Massachusetts 02452-6319, USA.

The many faces of Newton

Newton: The Making of a Genius

by Patricia Fara

Macmillan/Columbia University Press: 2002. 368 pp. £20/\$34.95

George Rousseau

This is not another biography of Isaac Newton. Instead, Patricia Fara, an experienced author on the history of early modern science, has reconstructed Newton's images — scientific, professional, personal, religious and domestic — for the light they shed on the public perception of great scientists. Her work provides a better understanding of this thoroughly paradoxical figure.

Fara examines Newton from many angles: as a symbol of genius and an icon of greatness; as viewed by his friends and enemies; as judged by his reputation in England and abroad; in view of the myths that developed about him and the shrines that were dedicated to him; and by glancing at a few of those scientists (Einstein, the Russian physicist and historian Boris Hessen and X-ray crystallographer J. D. Bernal) who are seen as his followers or heirs. The chapter on Newton's disciples in the eighteenth century epitomizes the way in which the complex transmission of newtonianism was achieved.

The separation of Newton's profiles into these categories magnifies the main theme of the book: that a coherent view of Newton is nowhere to be found. Only a painstaking biographer who chronicled every day, or every hour, of his life could make a case for the integration of character and behaviour in a figure as paradoxical as Newton. The task has been tried many times but always fails. The necessary documents do not survive, as they do for Charles Darwin, for example, and much is necessarily left to conjecture.

Only by demonstrating the impossibility of locating an essential Newton does Fara's strategy succeed. Psychobiography aims to integrate the components of individual lives by reducing all the strands to a few essential motives. The psychobiographer searches for the fundamental pulse of the subject and demonstrates how character and action are guided by it. Frank Manuel achieved it in his 1963 landmark psychobiography of Newton, *Isaac Newton, Historian* (Harvard University Press). Fara mentions this book only briefly, suggesting that Manuel's approach may be as problematic as all the others.

The crux of the problem is the difficulty of integrating Newton's incommensurate traits. Reading this book will persuade you that no core newtonian self

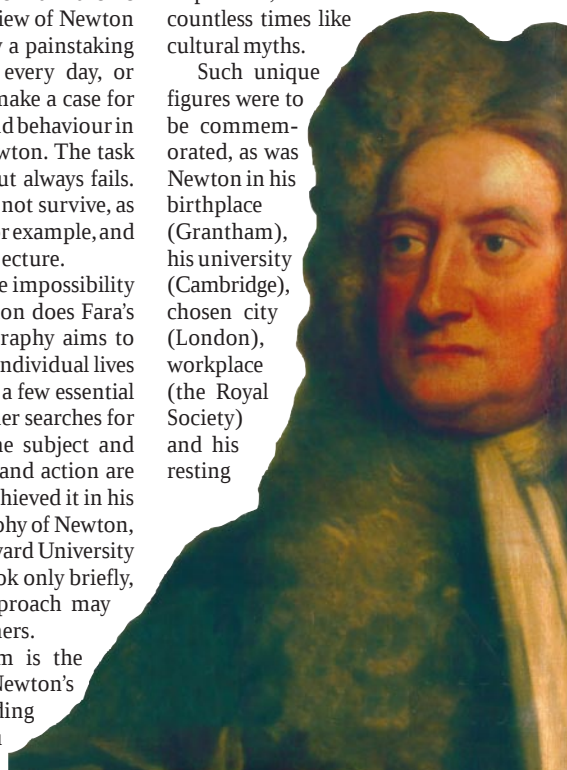
ever existed — or, if it did, that it cannot be retrieved. Only diverse views emphasizing scattered aspects of the man persisted, as was recognized during the Victorian era. The disparity is widespread: Newton was seen in different ways in France during and after the French Revolution, and in Russia and America, for example.

The final four chapters, dealing with myths that arose after 1800, are the most illuminating. The chapter on Newton as genius shows how he was configured by rationalists and by Romantics as a super-human creature. The difference lay in their emphasis. In around 1820 it was thought

that progress and perfectibility could be traced to the advances of a handful of geniuses who had discovered something fundamental (such as gravity, calculus and optics) that were destined to change the world. But individual genius still had to be explained.

This quest to celebrate genius was culturally sanctioned: society needed geniuses if it was to rid itself of calamity and to march towards a more civilized point. But the nineteenth-century cults of genius also gave rise to myths, such as the claim of Romantic poets William Blake and William Wordsworth that disembodied geniuses of Newton's calibre were terrifying. Genius was seen to be endowed, not earned — it was a divine gift permitting instant insight and magisterial discovery. The children's stories of Newton and his apple, and the tales about his extraordinary discoveries derived from inspiration, were retold countless times like cultural myths.

Such unique figures were to be commemorated, as was Newton in his birthplace (Grantham), his university (Cambridge), chosen city (London), workplace (the Royal Society) and his resting



place (Westminster Abbey). Yet those who opposed Newton must also be remembered: Charles Darwin refused to be buried beside Newton in Westminster Abbey; Einstein challenged Newton's physics; and John Maynard Keynes and his economics colleagues exclaimed, as Fara writes, that "Newton was not the first great scientist, but the last great magician".

Magic and science have never been easy bedfellows, despite pronouncements about their historical proximity in the early modern world. The images of magicians and scientists are different, and image-making within national settings is what this book is all about. Yet as time unfolds it becomes increasingly less important to enquire about the degree to which a particular scientist amounted to pure, disembodied genius, let alone to explore the quantum of his magic. We can assess Einstein's achievements, for example, without resorting to such extremes. If he was a genius, he was also an impoverished German Jewish immigrant who fled persecution to America. Nor do we worry whether he was unrivalled. The high-stakes, image-making process has been transferred from scientists to athletes, movie stars and politicians.

But there is one more turn of the screw. As Fara writes, "this duality as both an insane genius and a dispassionate scientist is unique to Newton". Unique it may be, but as time passes the possibility of an integrated image of Newton grows ever more remote, and in a century's time a unified view of Newton will appear even more elusive than it is now. This may not be so bad, or surprising: the lives of even the greatest scientists are no more rational or coherent than they are for the rest of us. They never were and never will be. ■

George Rousseau, former regius professor of English at King's College, Aberdeen, is research professor in the School of Humanities, De Montfort University, Leicester LE1 9BH, UK.

Untangling quantum mechanics

Entanglement: The Greatest Mystery in Physics

by Amir Aczel

Four Walls Eight Windows/John Wiley: 2002. 320 pp. \$25/£16.99

Vlatko Vedral

Popularizing physics is no easy task. The basic laws of quantum mechanics in particular contain many counterintuitive features that lay readers might find difficult to understand. So, in order to convey the message, there is often a tendency to oversimplify facts to the point of making them incorrect. Regrettably, this makes quantum physics much more mysterious than it really is.



Dead or alive? The state of the cashier is correlated with whether Lee Marvin's gun has been fired.

Amir Aczel falls prey to this when describing entanglement, a key feature of quantum mechanics, in his new book. As many other authors have done, he emphasizes that when two quantum systems are entangled, and then we measure one of the systems, we immediately know (or affect) the state of the other, no matter how far apart the systems are. We are led to believe that this is somehow weird.

However, in the day-to-day world that is well described by classical newtonian physics, we often encounter correlations. For example, imagine you are observing a bank robbery. The robber is pointing a gun at the terrified bank clerk. By looking at the clerk you can tell whether the gun has gone off or not. If the clerk is alive and unharmed, you know that the gun has not fired. If the clerk is lying dead, you know that the gun has fired. There is therefore a direct correlation between the state of the gun and the state of the clerk. 'Gun fired' means 'clerk dead', and 'gun not fired' means 'clerk alive'. Of course, we assume that the robber shoots to kill and never misses.

The key (and surprising) observation is that quantum-mechanical systems may be further correlated such that the gun is in the superposition 'fired and not fired', and the clerk is then in the state 'dead and alive' at the same time! Quantum mechanically, as a consequence of this superposition, there is simply more correlation between atoms or photons than we would expect classically. This kind of quantum 'supercorrelation', first quantified precisely by John Bell in 1964, is what we refer to as 'entanglement'.

In the early days of quantum mechanics there were several attempts to use entanglement to uncover a paradox in the foundations of quantum physics. However, by quantifying these correlations, Bell took

them out of the philosophical realm of Einstein and Niels Bohr, and demonstrated how entanglement and the completeness of quantum mechanics could coexist. These extra correlations are 'real', in that they have been confirmed by experiment and, more importantly, have been successfully applied to quantum teleportation (which would otherwise be impossible) and to quantum cryptography.

The oversimplification of entanglement notwithstanding, I find that Aczel has done a good job of describing its colourful history, as well as modern developments such as teleportation. But what he perhaps does best is enlighten us with some snippets about the key physicists involved. For example, I found many interesting and humorous anecdotes about Erwin Schrödinger's love life, about John von Neumann being perceived as an alien by the US public, and about the friendship between Heisenberg and Bohr. There is, however, a dangerous tendency in popular science to idolize scientists beyond justification, making them appear more extraordinary than they really are. This book walks a fine line between historical accuracy and fictional infatuation.

I am altogether happy that there is now finally a book on entanglement, almost 70 years after its discovery, and recommend it to people interested in the historical background and practical implications of quantum mechanics. I am afraid, however, that anyone interested in understanding quantum mechanics, and the way in which scientists think about it, is much better off reading any of Richard Feynman's popular accounts, such as *QED: The Strange Theory of Light and Matter* (Princeton University Press, 1985). ■

Vlatko Vedral is at the Blackett Laboratory, Imperial College, Prince Consort Road, London SW7 2BZ, UK.

Intermediate impact factors

Robert Jedicke

No statistical documentation of objects of a certain size that enter Earth's atmosphere has hitherto been available. Analysis of data from US government satellites has bridged the gap.

The natural world that scientists explore often seems pathologically arranged to thwart discovery. To peer inside the nucleus, we must build massive accelerators and detectors the size of office buildings; the mysterious beginnings of our Universe are shrouded by tremendous distances, requiring telescope mirrors as big as swimming pools to collect photons liberated billions of years ago. So there is something satisfying about turning nature to our own purposes, applying its peculiar properties to explore the edges of our knowledge and forcing nature itself to become the tool with which we explore the Universe. On page 294 of this issue, Brown *et al.*¹ describe how they have used a detector as big as the Earth itself — the planet's protective, blanketing atmosphere, which acts as a scintillator to detect the powerful but rare impacts of small asteroids or comets.

Meteors, those romance-inspiring flashes of light in the night sky, are created when small pieces of dirt called meteoroids slam into the Earth's atmosphere at speeds of tens of kilometres per second (Fig. 1). These fragments are the detritus of collisions between asteroids in the belt between Mars and Jupiter, or debris spawned from cometary nuclei, and their orbits gradually evolve onto paths that cause them to collide with the Earth. The burnt remains of these objects rain down on Earth and increase its mass by many tonnes every day — although the brightest meteor you may ever see might be the death beacon of a rock no bigger than a pea. As they are small and plentiful, it isn't particularly difficult to study the statistics of these impacts. They are found and measured regularly by meteor radar facilities² and automated wide-field camera searches³, and even by examining the surfaces of objects that have been in orbit about the Earth⁴.

Tipping the scale in the other direction, massive mountains of rock and ice (asteroids and comets, respectively, in planetary jargon, though the distinction probably isn't as strict as was once thought), which may be many kilometres in diameter, also move in orbits that may one day intercept the Earth. The latest estimates for the number of these objects larger than a kilometre in diameter is about a thousand⁵, and there is a good chance that one of them will strike the Earth within the next million years. Theoretical predictions⁶ indicate that when objects this



Figure 1 Shooting stars. The Leonid meteor shower, which happens each year between 17 and 19 November, can be a breathtaking sight: small rocks, fragments of the comet Tempel–Tuttle, burn up or even explode as they hit the Earth's atmosphere. This meteor (inset) was photographed in the spectacular shower of 1966. Fortunately, impacts of much larger objects are rare. In 1908, a meteor as much as 50 m in diameter exploded in the Earth's atmosphere over Tunguska, Siberia. The resulting shock wave flattened trees in the surrounding area over hundreds of square kilometres (main image). Using satellite data, Brown *et al.*¹ have calculated the average rate of such devastating impacts to be about one per thousand years.

big collide with the Earth, there will be severe environmental consequences and massive continental-scale destruction. Fortunately for us, because these are relatively big objects, and because they are in orbits that swing them through the Earth's 'backyard' on occasion, the chances are good that planetary astronomers will discover the next big one before a collision takes place. Several groups around the world are surveying the night skies for these objects using modest-sized (about 1-m diameter) telescopes, and they are currently about halfway through cataloguing all of these most dangerous objects.

In between the harmless little meteors and the civilization-ending flying mountains lies an intermediate regime of objects,

ranging from one metre to tens of metres in diameter. These objects are too small to be detected with telescopes and too uncommon in their rate of impact to be regularly visible from the planet's surface as they ignite in the upper atmosphere. Early extrapolations⁷ of the number distribution from the smallest to the largest size scales seemed to indicate a discontinuity (or more than one) between the two extremes — a condition ripe with opportunity for a scientific Goldilocks to identify the detector that is 'just right' for the task.

The detector in this case is the Earth's atmosphere. When a large meteoroid ploughs into its rarefied upper reaches, shock waves catastrophically disrupt the object, creating a blast equivalent to the

explosion of many kilotonnes of TNT. Some of this energy is converted into visible light and other forms of radiation that can be detected by instruments on satellites operated by the US Departments of Defense and Energy, and these data have been used by Brown and colleagues¹. Although there are numerous uncertainties involved in correcting the observed signal intensity at the satellite to obtain the incoming energy and the size of the impactor, Brown *et al.* have succeeded in documenting this previously unexplored region of the size or energy distribution of objects in the neighbourhood of the Earth.

The most satisfying aspect of their measurement is the fact that the bridge between the larger and smaller objects fits the gap so well. It's as though a bridge were designed in Australia, built in France, transported to the United States and dropped seamlessly into a location in the Grand Canyon. Indeed, to within the errors of measurement, a single power-law seems to fit the size or energy distribution over ten orders of magnitude in energy (see Fig. 4 on page 296). This simple behaviour has been predicted analytically for self-similar collisional cascades⁸, in which a set of objects whose physical strength is independent of their size grind into, strike and disrupt one another.

In 1908 a tremendous explosion, estimated to be equivalent to about ten megatonnes of TNT, flattened trees over hundreds of square kilometres near Tunguska, Siberia (Fig. 1). If a similar event were to take place over a densely populated region of the world today, the death toll could easily be many millions of people. By extending the bridge they had measured by about 2.5 orders of magnitude in energy, Brown *et al.* estimate that the average time between impacts of this energy is about a thousand years — five times longer than was thought only ten or twenty years ago. Their results for a Tunguska-like impactor dovetail nicely with earlier and unpublished measurements by others, as shown in their Fig. 4. Because there are fewer of these objects than originally believed, we can all worry a little less about the risk of the next hazardous impact. Yet it is essential to emphasize that the job of cataloguing all objects down to the size of the Tunguska parent body is not currently achievable, and probably will not be completed for many decades.

The timely measurement by Brown *et al.*¹ of the size or energy distribution of objects in this range has linked the fields of meteor and comet/asteroid planetary astronomy in a manner that shows they are not merely distant relatives but kissing cousins. It seems that, as well as recognizing the Earth's atmosphere as a 'just right' detector system, Goldilocks has charmed the US government into releasing valuable scientific information obtained serendipitously while

its satellites keep watch over the world's nuclear arsenals.

Robert Jedicke is in the Lunar and Planetary Laboratory, University of Arizona, 1629 E University Boulevard, Tucson, Arizona 85721-0092, USA.

e-mail: jedicke@pirlmail.lpl.arizona.edu

1. Brown, P., Spalding, R. E., ReVelle, D. O., Tagliaferri, E. & Worden, S. P. *Nature* **420**, 294–296 (2002).

2. Cepelch, Z. *et al. Space Sci. Rev.* **84**, 327–471 (1998).
3. Stokes, G. H. *et al. Icarus* **148**, 21–28 (2000).
4. Lurance, M. R. & Brownlee, D. E. *Nature* **323**, 136–138 (1986).
5. Stuart, J. S. *Science* **294**, 1691–1693 (2001).
6. Lewis, J. S. *Comet and Asteroid Impact Hazards on a Populated Earth: Computer Modeling* (Academic, San Diego, 2000).
7. Rabinowitz, D. L. *Astrophys. J.* **407**, 412–427 (1993).
8. Williams, D. R. & Wetherill, G. W. *Icarus* **107**, 117–128 (1994).

Developmental biology

Colon construction

Mark Peifer

It is no mean feat for organisms to make and maintain their organs. The complex cellular and molecular processes involved are illustrated by two studies of the proteins that participate in producing a colon.

We often forget what marvellous machines our bodies are. Each of our organs is constructed from diverse cell types arranged in a complex, stereotyped pattern, allowing them to carry out their assigned tasks — propelling blood, composing a paragraph, or absorbing nutrients. Perhaps more remarkably, these organs operate continuously for decades, requiring constant remodelling to replace cells lost to attrition. In two landmark papers in *Cell*^{1,2}, Clevers and colleagues begin to explain the structure of one organ, the intestine, identifying the architect that directs its development and renewal, contractors that supervise different aspects of the process, and skilled labourers that do the heavy lifting. In doing so they touch on many of the hottest issues in biology, including stem cells and microarrays, as well as perennial favourites such as tumour suppressors and cell-cycle regulators.

The intestine is an interface with the outside world, serving to protect the body and absorb nutrients. During development this organ must produce an array of cell types with different roles and must position each cell properly — a task that is complicated by the harsh environment inside the intestine. As a result, cells live only three to five days, so throughout life the intestine must constantly recreate its complex architecture, with new cells taking the place of their deceased predecessors. Intestinal cells are arranged in a folded sheet known as an epithelium (Fig. 1; reviewed in ref. 3). The bases of the folds, the crypts, contain stem cells and their transiently proliferating daughters. As the cells so produced become specialized (differentiate), most migrate upward along the lengths of the folds — the villi — to form various cell types that absorb nutrients and perform the epithelium's other functions. So-called Paneth cells, meanwhile, migrate downward to the crypt base.

How is this complex architecture pro-

duced and maintained? Over the years, beginning with genetic analyses in model animals, scientists have identified and examined the function of key molecules that direct developmental decisions, revealing an outline of how cells throughout the embryo choose their fates. As this work has matured, it has become clear — often through the discovery of unexpected connections between developmental regulators and cancer — that the machinery used to establish tissues is also used to maintain them in adults.

One superb example involves the Wnt proteins (reviewed in ref. 4), extracellular signalling molecules that are key to the patterning of diverse tissues in fly and mammalian development. Wnt signals are transmitted into cells by a signal-transduction pathway that ultimately activates gene-expression programmes. This occurs through inactivation of a multiprotein complex, including the tumour suppressor APC, that otherwise targets another protein, β -catenin, for destruction. Inactivation of this complex stabilizes β -catenin, which can then enter the nucleus, where it converts DNA-binding proteins of the TCF/LEF family from gene repressors to gene activators.

Inappropriate activation of Wnt signalling underlies most cases of colon cancer (reviewed in ref. 5). For instance, loss of APC leads to the formation of benign polyps, which become malignant tumours if they undergo additional genetic changes. Without APC, β -catenin is stabilized as described above, and Wnt target genes are activated, including the *Myc* and *cyclin D1* genes^{6–8}, which drive cell proliferation and are thought to help drive tumour formation. These findings suggested that normal Wnt signalling might regulate normal colon cell proliferation. Supporting that idea, mice lacking the DNA-binding protein TCF-4 die soon after birth, and their colons lack crypts, as they fail to maintain intestinal cell proliferation⁹.

To explore this idea further, the Clevers

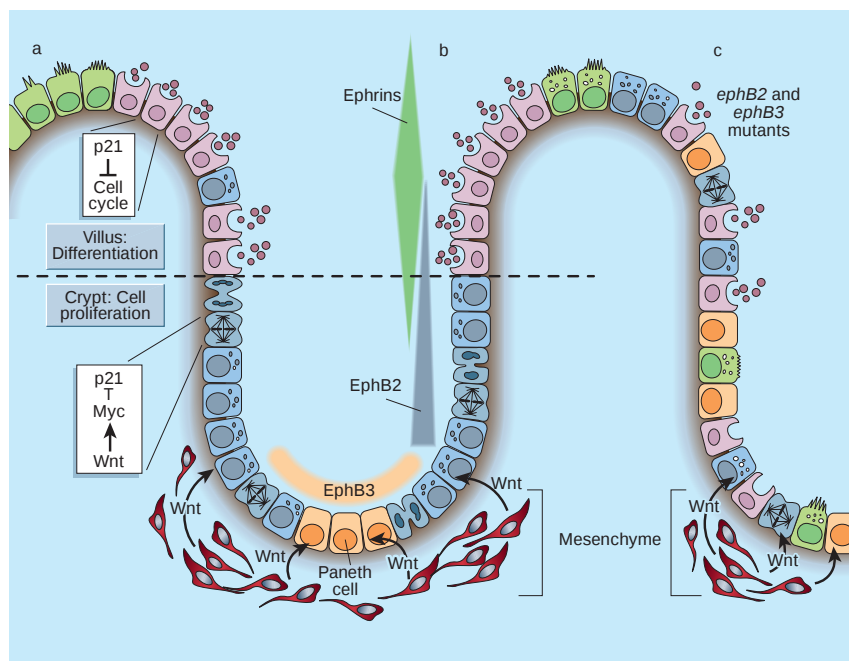


Figure 1 The architects, contractors and labourers in colon construction. The intestine is lined by a folded epithelial sheet. Proliferating cells are confined to the crypts; most of the differentiated cells are in villi, although Paneth cells are found at the crypt base. **a**, Clevers and colleagues^{1,2} suggest that Wnt proteins, presumably released from mesenchymal cells beneath the crypt, maintain the proliferative potential of crypt cells by turning on the gene encoding the Myc protein, which in turn represses the cell-cycle inhibitor p21. In villi, no Wnt signal is received and p21 mediates cell-cycle arrest, thus allowing differentiation. **b**, The authors also find that ephrin proteins are expressed on cells of villi and at the top of crypts, whereas their receptors, Eph proteins, are expressed in overlapping patterns in the crypt. The authors thus propose that repulsive interactions between ephrins and Eph receptors allow proliferating and differentiating cells to separate out. **c**, This idea is supported by the fact that in mice with mutations in *ephB2* and *ephB3*, cells are no longer positioned correctly. (Figure derived from one kindly provided by E. Sancho and H. Clevers.)

lab began¹ by studying proliferating colon cell lines carrying mutant versions of β -catenin or APC, which constitutively activate Wnt signalling. When the authors turned off Wnt signalling in these cells, the cells ceased dividing, consistent with the notion that Wnts are indeed needed for proliferation. Next, Clevers and colleagues looked for regulators of the cell-division cycle that could mediate this switch. Their data point to the protein p21, which inhibits certain enzymes that drive cell-cycle progression. Wnt signalling downregulates p21 expression (Fig. 1a), but when such signalling is blocked, p21 is expressed and causes cell-cycle arrest. In an interesting aside, blocking the cell cycle by overexpressing p21 induces differentiation, suggesting an intriguing coupling between cell division and differentiation. Thus, p21 is the labourer that shapes the cell cycle.

The authors next investigated whether the TCF- β -catenin complex supervises p21 directly, or whether it farms out this task to a sub-contractor. The latter appears to be the case. Wnt signals turn on *Myc*, which encodes a gene-transcription factor. *Myc* then mediates the proliferative effects of Wnts, apparently by directly repressing p21 expression.

Clevers and colleagues then looked for other effects of Wnts on colon cells, using microarrays to examine the expression of some 24,000 genes in the presence or absence of Wnt signalling¹. The response was surprisingly simple — 120 genes were activated by Wnt signalling and 115 were repressed. They next tested the hypothesis that Wnt signalling promotes cell proliferation in colon crypts. This was confirmed in striking fashion. Genes activated by Wnts were expressed in the proliferative compartment of crypts (and are overexpressed in colorectal tumours) (Fig. 1a). Conversely, genes downregulated by Wnts were expressed in specialized cells at the tops of crypts and in villi. Together, these data suggest that Wnts are colon architects, drafting plans that direct the switch from proliferation to differentiation. Wnt signals, presumably emanating from mesenchymal cells beneath the crypt, maintain crypt cells in a proliferative state. Cells exiting this environment are no longer exposed to Wnts, and differentiate rather than proliferate. But how do cells position themselves properly in this complex epithelium? The Clevers group aimed to find out.

The success of a microarray experiment depends on creativity in mining surprising



100 YEARS AGO

Some very interesting observations relative to the cause and nature of radio-activity have been recently made by Messrs. Rutherford and Soddy, an account of which is given in the September number of the *Philosophical Magazine*. The experiments were carried out with thorium compounds, all of which are radio-active. The authors arrive at the conclusion that the greater part of the radio-activity of thorium is due to a non-thorium type of matter, represented symbolically by ThX, possessing distinct chemical properties. The activity of this new type is not permanent, but undergoes a gradual process of decay, the value falling to one half in about four days. The constant radio-activity of thorium is supposed to be maintained by the continuous production of this new type of matter from the thorium compounds. Its rate of production and the rate of decay of its activity appear to be independent of the physical and chemical conditions of the system. The ThX is capable of exciting radio-activity on surrounding inactive bodies, and about 20 per cent. of the total activity of thorium is due to this action of the ThX. By suitable means, thorium can be freed from both ThX and the excited radio-activity produced by the latter, and then possesses an activity about 25 per cent. of its original value. This latter, the authors believe, is due to a second non-thorium type of matter.

From *Nature* 20 November 1902.

50 YEARS AGO

Fifty years ago, in 1902, Starling, working in University College, London, suddenly had a new idea that [pancreatic secretion] must be a chemical reflex, and quickly proved this... Starling and his fellow-worker, Bayliss, postulated that the higher animals possess, in addition to the nervous system, a second co-ordinating and integrating system: substances formed in one organ circulate in the blood and regulate the function of another organ or site. To these they gave the name 'hormones' (to excite), and said that, when this word was used, the attributes of messenger should be understood. This idea of a blood-borne regulation of normal body functions has proved to be one of the most fruitful in physiology, and to-day, with the advent of cortisone and adrenocorticotrophic hormone, may be ready to give us a new concept of disease.

From *Nature* 22 November 1952.

gems from gene lists. Clevers and colleagues noticed that the receptor proteins EphB2 and EphB3 are upregulated by Wnt signalling, whereas their ligand, ephrin-B1, is down-regulated². These proteins have properties (reviewed in ref. 10) that caught the eye of the investigators. Unlike many ligands, the ephrins remain tethered to the cell that makes them. Thus, both ligands and receptors are found on cell membranes, and interactions between them often lead to cell repulsion, so that cells expressing ligand on their surface are sorted out from those expressing receptor. This helps set segmental boundaries in the brain and shapes retinal axon guidance, for example.

Eph receptors are expressed in overlapping regions of crypts, consistent with their activation by Wnt signalling, and ephrins are expressed in a complementary domain in differentiated cells (Fig. 1b). Clevers and colleagues² hypothesized that repulsive Eph–ephrin interactions might establish a boundary between proliferating and differentiated cells, and thus position cells properly. To test this, they turned to mice with mutations in EphB2 and EphB3, and found that many different cell types were mispositioned all along the crypt–villus axis (Fig. 1c). Proliferative cells were no longer restricted to crypts; differentiated cell types strayed from villi; and Paneth cells left the crypt base.

These data firmly establish Wnt signals as colon architects, identify sub-contractors such as Myc that supervise particular parts of the job, and highlight skilled labourers, such as p21 and Eph receptors, that carry out

specialized tasks. They also raise new questions. How much of the programme is run directly by TCF– β -catenin and how much is parcelled out to sub-contractors such as Myc? How are proliferation and differentiation coupled? Perhaps most intriguing, how do intestinal stem cells generate progeny with asymmetries in developmental potential, and what are the signals or intrinsic factors that cooperate with Wnts in this process? The known role of Wnt signalling in asymmetric cell divisions in other contexts is intriguing in this regard. Finally, do other tissues use similar mechanisms? Studies of the skin hint that this may be the case. Wnts, presumably emanating from mesenchymal cells at the base of the hair follicle, are key to cell-fate decisions there (reviewed in ref. 11) — it will be interesting to see whether these proteins regulate skin-cell positioning, and whether Eph receptors are involved. ■

Mark Peifer is in the Department of Biology and the Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3280, USA. e-mail: peifer@unc.edu

1. van de Wetering, M. *et al.* *Cell* **111**, 241–250 (2002).
2. Batlle, E. *et al.* *Cell* **111**, 251–263 (2002).
3. Stappenbeck, T. S., Wong, M. H., Saam, J. R., Mysorekar, I. U. & Gordon, J. I. *Curr. Opin. Cell Biol.* **10**, 702–709 (1998).
4. Cadigan, K. M. & Nusse, R. *Genes Dev.* **11**, 3286–3305 (1997).
5. Peifer, M. & Polakis, P. *Science* **287**, 1606–1609 (2000).
6. He, T. C. *et al.* *Science* **281**, 1509–1512 (1998).
7. Tetsu, O. & McCormick, F. *Nature* **398**, 422–426 (1999).
8. Shutman, M. *et al.* *Proc. Natl Acad. Sci. USA* **96**, 5522–5527 (1999).
9. Korinek, V. *et al.* *Nature Genet.* **19**, 379–383 (1998).
10. Wilkinson, D. G. *Nature Rev. Neurosci.* **2**, 155–164 (2001).
11. Fuchs, E., Merrill, B. J., Jamora, C. & DasGupta, R. *Dev. Cell* **1**, 13–25 (2001).

Materials chemistry

Crystals in a flash

David W. Oxtoby

If laser light is shone on a solution, the crystal structure that forms depends on the light polarization, and the more intense the laser, the greater the probability of crystal nucleation. The challenge now is to work out why this is.

Molecules can crystallize into a variety of structures, known as polymorphs. A particular structure can sometimes be selected by varying the solvent from which the crystal is grown or by adding a small concentration of an inhibitor¹, but this is inconvenient and slow. In *Physical Review Letters*, Garetz *et al.*² show that structure selection can be made literally in a flash of light, by irradiating a supersaturated solution with laser light of a particular polarization. Circularly polarized light causes the α -phase of crystalline glycine to develop from aqueous solution, whereas the γ -phase results from illumination with linearly polarized light. Their result suggests a new method for controlling polymorph selection

that could be of use in the pharmaceutical industry, and which also provides insight into the detailed mechanism of crystal nucleation, the first step in the formation of crystals from solution.

Several years ago, Garetz *et al.*³ showed that crystal nucleation from solution can be speeded up by laser irradiation. In that experiment, supersaturated solutions of urea, which took days or weeks to crystallize in the absence of laser light, nucleated during a single, nine-nanosecond laser pulse — an acceleration of the nucleation process by 13 orders of magnitude. Even in the nineteenth century, photochemical nucleation was a well-known phenomenon, in which absorption of light causes ionization or other chem-

ical reactions that lead to the formation of a new phase. But the authors argued³ that what they were seeing was not a photochemical process. Their new experiment² provides conclusive evidence for this, because a photochemical reaction would not be affected by different laser polarizations.

Instead, Garetz *et al.* attribute the effect to a molecule's polarizability, which is acted on by the electric field of the laser light. A small dipole moment is induced and the molecule's energy is lowered. This favours a molecular orientation in which the largest component of the polarizability lies along the direction of polarization of the laser field. For isolated molecules the effect is small and is overwhelmed by thermal fluctuations. But a supersaturated solution contains clusters of molecules that have not yet become fully crystalline and can be aligned by appropriate laser fields, moving them along the path to crystallization.

From their examination of the crystal structures of α - and γ -glycine, Garetz *et al.*² argue that rod-like clusters of glycine molecules in water — which are acted on more strongly by linearly polarized light — are the precursors for γ -glycine, which consists of helical chains of glycine molecules arranged head to tail. On the other hand, disk-like clusters in solution should be more strongly stabilized by circularly polarized light, and are the precursors for α -glycine. These qualitative predictions are consistent with the new experimental observations.

Garetz *et al.* also studied the quantitative effect of laser intensity on crystallization, and found that the probability of nucleation increased with intensity in a monotonic, though nonlinear, fashion. (For experimental reasons, to trace the relationship, the authors returned to using urea solutions in place of glycine.) Even at the highest power levels used, however, fewer than half of the samples (of either urea or glycine) crystallized. It is not clear why the other samples did not nucleate. Although nucleation is a random process, it should be reproducible on average. Further study of this is clearly needed.

The magnitude of the laser effect is also not yet understood. Quantitative estimates of the lowering of the energy barrier to nucleation seem insufficient to account for the results, and the authors suggest that the laser fields may instead affect the 'prefactor' — a factor that includes the kinematics of the system in the calculation of the nucleation rate. But this is also unlikely to be a large effect, because prefactors rarely change by more than one or at most a few orders of magnitude when the physical conditions change. Perhaps a cooperative effect is taking place, in which a number of small-amplitude perturbations change the entire structure of a cluster and lower its free energy significantly. This question deserves investigation, and Garetz *et al.*² have laid down a challenge to

theorists in their paper, pointing out that the strong interactions in glycine clusters are beyond the reach of present computational capabilities.

The effects of an electric field on crystal nucleation have already been studied for a different problem — that of simulating ice formation in supercooled water. In the absence of an external field it is extremely difficult to get ice to freeze in a computer simulation, and in fact this has been achieved only once⁴, after heroic and lengthy efforts. In contrast, when water crystallization is simulated in a static external electric field, ice nucleates readily⁵. Perhaps this process is analogous to glycine crystallization, in that partially oriented water clusters might pass more easily from a liquid-like to a crystalline structure, initiating crystallization of the

entire sample. The presence of a local electric field may also be critical in nucleating crystals at the interface between a solution and an amphiphilic monolayer⁶. These hypotheses remain to be explored, through both theory and experiment. ■

David W. Oxtoby is at the James Franck Institute and the Department of Chemistry, University of Chicago, Chicago, Illinois 60637, USA.
e-mail: d-oxtoby@uchicago.edu

1. Weissbuch, I., Leiserowitz, L. & Lahav, M. *Adv. Mater.* **6**, 953–966 (1994).
2. Garetz, B. A., Matic, J. & Myerson, A. S. *Phys. Rev. Lett.* **89**, 175501 (2002).
3. Garetz, B. A., Aber, J. E., Goddard, N. L., Young, R. G. & Myerson, A. S. *Phys. Rev. Lett.* **77**, 3475–3478 (1996).
4. Matsumoto, M., Saito, S. & Ohmine, I. *Nature* **416**, 409–413 (2002).
5. Svishchev, I. M. & Kusalik, P. G. *Phys. Rev. Lett.* **73**, 975–978 (1994).
6. Rapaport, H. *et al. J. Phys. Chem.* **104**, 1399–1428 (2000).

Molecular evolution

Booting up life

Gerald F. Joyce

A key question about evolution is how the first informational molecules — thought to be an early form of life — could generate efficient self-replication machinery. The problem is tackled in new computer simulations.

In thinking about the origin of life on Earth and, in principle, elsewhere in the Universe, many regard the ‘hard problem’ to be getting the ball rolling in the first place. It is generally thought that life on Earth began with the production of macromolecules that served as primitive stores of genetic information (genomes). Crucially, there was a need to replicate these macromolecules, both to propagate the contained information and to allow the occurrence of mutations — which, within limits, are essential for darwinian evolution based on natural selection. The rest would then follow naturally: darwinian innovation would provide the solutions to challenges imposed by a changing environment.

But the need for replication brings up another hard problem: it is not clear whether, in life’s earliest stages, genomes were large enough to encode an efficient and reasonably accurate machinery for copying themselves, as is the case today. The crux of the matter is that a sophisticated replication machinery is needed to maintain a sizeable genome, but a sizeable genome is needed in the first place to encode a sophisticated replication machinery. This is a difficulty that cannot be waved away by assuming a bootstrapping process of steady improvement in replication efficiency and fidelity, permitting progressively larger genomes. On page 340 of this issue¹, Szabó *et al.* address this problem through computer simulations of evolutionary bootstrapping. They find that it is possible to evolve towards larger

genomes that encode better replication machineries — but only if some form of spatial isolation is imposed on the population of replicating genomes.

As first discussed by Manfred Eigen more than 30 years ago², there is likely to have been an information crisis during the earliest stages of life. Eigen described an error threshold that relates the rate and fidelity of replication of individuals in a population to the maximum genome size that can be maintained over successive generations. Genome size is especially sensitive to replication fidelity, with the maximum number of residues in the genome being roughly inversely proportional to the replication error rate per residue.

So, for a simple genome of 100 nucleo-

tides, the error rate of replication must be no more than about 1% per nucleotide. Biological systems tend to operate close to the brink of the error threshold: bacterial viruses with a genome of some 10⁴ residues typically have an error rate of 10^{−4} per residue; for bacteria with a genome of about 10⁶ residues, the rate is up to 10^{−6}; and multicellular organisms with a genome of some 10⁹ residues have an error rate of up to 10^{−9}. By living close to the edge, evolving systems can explore the broadest possible range of errors while still retaining information that has proved useful over time.

Before the discovery of catalytic RNA molecules (ribozymes), the existence of the error threshold posed a formidable challenge for those seeking to understand how darwinian evolution began. If life at the outset, as now, was based on genomes consisting of nucleic acids and a replication machinery comprising proteins, then the genome must have been sufficiently complex to encode those proteins as well as the machinery for protein synthesis. Such a large genome would have required a highly accurate replication machinery, allowing it to operate within the error threshold.

If the genome and ‘replicase’ were embodied in the same molecule, however, then a much smaller genome would have been required. Indeed, it is now widely believed that an ancestral form of life on Earth was based on RNA genomes that were copied by an RNA replicase^{3,4}. No such replicase has been described, but a ribozyme has been produced in the laboratory that contains about 200 nucleotides and synthesizes RNA molecules of up to 14 nucleotides, with an error rate of about 3% per residue⁵. So one could imagine a small RNA genome that operates as an RNA replicase with modest efficiency and fidelity. Through darwinian evolution this could give rise to a succession of improved replicases and ever larger genomes.

Szabó *et al.*¹ investigated whether this could actually happen through their computer simulations. They constructed a digital

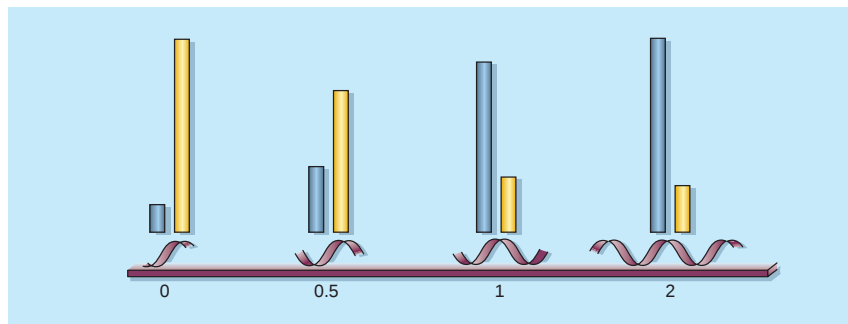


Figure 1 Genome evolution in a computer. Szabó *et al.*¹ created a digital population of genomes that were immobilized on a surface and found that, over the course of two million steps of computer-simulated evolution, the average genome length (purple) increased more than sixfold, while replicase activity (blue) increased more than sevenfold and replicase error rate (yellow) decreased more than fourfold. The average behaviour for the population is shown after 0, 0.5, 1 and 2 million steps.

population of evolving genomes, each of which served as both a template for its own replication and a catalyst for the replication of other genomes. Different sequences within a genome were used to encode the traits of template efficiency, replicase activity and replicase fidelity. Accumulation of sequence information for a particular trait led to greater expression of that trait, subject to diminishing returns. The accumulating sequence, however, required ever more efficient replication. This could come about through mutation, which resulted in changes in the length or sequence of a genome, and hence changes in its corresponding behaviour. Thus the simulation tested whether darwinian evolution could indeed bring about a bootstrapping of increased genome size and improved replicase function. Although inspired by thinking about RNA-replicating ribozymes, the simulation should be regarded as applying generically to the evolution of genome replicases.

A crucial variable in the simulation turned out to be the degree of spatial isolation imposed on the evolving population. The simulated molecules inhabited a two-dimensional grid, and could either replicate, or be replicated by, their neighbours. If they

were allowed to diffuse rapidly across the grid, then the population became dominated by selfish molecules with small genomes that were copied readily by other replicases, but encoded only rudimentary replicase activity themselves. Plagued by these parasites, the population as a whole could not evolve sizeable genomes and robust replicases. If, however, diffusion was restricted to roughly the same timescale as the rate of replication, then the outcome was completely different (Fig. 1). The grid became populated with local enclaves of large genomes encoding efficient replicases that brought about each other's replication. Some parasites were present, but these could not drag down the reciprocal altruists.

Most biologists are aware of the value of spatial isolation for organisms to develop selfish traits. Why, for instance, share the fruits of your metabolic labour with the surrounding environment when they can be sequestered for your own use within a cellular boundary or some other compartment? It is fascinating to realize that spatial isolation also contributes to good citizenship, at least in the world of molecules. Of course, one should not take the two-dimensional grid used by Szabó *et al.* too literally. This was use-

ful for conducting the simulation and may be viewed as a model for replicases that are confined to a surface^{6,7}. There are other ways, however, of achieving spatial isolation; for instance, molecules could be sequestered within a lipid vesicle or a non-covalent aggregate in solution. Spatial isolation provides an opportunity for darwinian evolution to 'discover' solutions to problems that require molecules to interact. The results obtained by Szabó *et al.* show that this makes evolutionary bootstrapping possible—and perhaps it is what started our planet on the path to the complex biosphere that exists today. ■

Gerald F. Joyce is in the Departments of Chemistry and Molecular Biology, and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.

e-mail: gjoyce@scripps.edu

1. Szabó, P., Scheuring, I., Czárán, T. & Szathmáry, E. *Nature* **420**, 340–343 (2002).
2. Eigen, M. *Naturwissenschaften* **58**, 465–523 (1971).
3. Gilbert, W. *Nature* **319**, 618 (1986).
4. Joyce, G. F. *Nature* **418**, 214–221 (2002).
5. Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E. & Bartel, D. P. *Science* **292**, 1319–1325 (2001).
6. Orgel, L. E. *Origins Life Evol. Biosphere* **28**, 227–234 (1998).
7. Luther, A., Brandsch, R. & von Kiedrowski, G. *Nature* **396**, 245–248 (1998).

Cancer

Super p53

One of the most intensively studied of cellular molecules, the p53 protein is well known for its role as a tumour suppressor: it helps to protect against cancer by regulating the response of our cells to damage and stress. A common characteristic of human cancers is the mutation, and hence inactivation, of p53, which predisposes organisms to tumours. Recent evidence suggests that the opposite may also be true — having more functional p53 than normal might prevent cancer.

Writing in *EMBO Journal* (**21**, 6225–6235; 2002), Isabel García-Cao *et al.* provide further support for this idea. The authors have generated mice that carry an extra, fully functional copy of the p53 gene, as seen in the image here, which shows the two normal (wild-type) p53 genes plus the extra one (the transgene) as bright dots on the blue-coloured chromosomes. These 'super-p53 mice' display an enhanced response to DNA damage. Remarkably, the animals are also less susceptible to spontaneous or carcinogen-induced tumour formation than their wild-type

counterparts. These results are consistent with a report earlier this year by Stuart D. Tyner *et al.* (*Nature* **415**, 45–53; 2002), which described mice carrying a truncated form of p53 that appeared to increase the activity of wild-type p53. Like the super-p53 mice, these animals were substantially protected against cancer.

Unexpectedly, however, the study by Tyner *et al.* suggested that preventing cancer by augmenting p53 activity may come at a price: mice carrying the truncated p53 mutant succumbed to premature ageing. This hinted that the mechanism used by p53 to protect against tumours may at the same time advance the ageing process. But in an interesting twist, García-Cao *et al.* now report that super-p53 mice have a normal lifespan, and show no evidence of early ageing.

How can this difference be explained? One possibility is that in the mice expressing truncated p53, p53 activity is constitutively enhanced, which might accelerate ageing. In contrast, the extra copy of



p53 in the super-p53 mice is subject to normal regulation, and basal levels of p53 activity are unaltered, suggesting that p53 activity is increased only when cells encounter stress.

Together, these studies offer the exciting prospect that using drugs to augment tumour-suppressor

activity could help in cancer prevention. But the differences between the two results serve as a reminder: we need to learn a lot more about tumour-suppressor function and regulation to safeguard against any potentially harmful side-effects of this approach.

Barbara Marte

Earth science

Ancient mantle in a modern plume

Elisabeth Widom

Osmium isotopes record evidence for 2.5-billion-year-old mantle beneath the Azores. The origin of this ancient mantle has implications for the nature and timescale of mantle convection.

The evolution of the Earth's interior is largely controlled by plate-tectonic processes. Convection of the mantle delivers hot material to the surface from depth — both at mid-ocean ridges, where shallow mantle melts to form new basaltic oceanic crust, and beneath ocean islands such as Hawaii, where rising mantle plumes melt to produce hotspot volcanism. The complementary process is subduction at convergent plate boundaries. Here, lithospheric plates — composed of oceanic crust and associated mantle depleted in the components that formed the crust — enter the Earth's interior.

The subducted lithosphere ultimately returns to the Earth's surface as a component of mantle plumes. But what is the nature and timescale of this recycling process? The results of Schaefer *et al.*, reported on page 304 of this issue¹, add a new dimension to our thinking about these questions. They provide the first unequivocal evidence for the presence of ancient lithospheric mantle in a plume source. If this is indeed recycled oceanic lithosphere, the implication is that it has been stored for billions of years deep within the Earth.

Long-term differences in the concentrations of trace elements in mantle and crustal reservoirs have led to present-day variations in their isotopic compositions, because radioactive parent isotopes decay and produce radiogenic daughters. The mantle that feeds mid-ocean-ridge systems has sustained a long-term depletion in incompatible elements (those that prefer to partition into melt) due to melting and extraction of crust, and thus has daughter-element isotopic compositions that are distinct from those of old crust enriched in incompatible elements. Isotopic differences between the shallow mantle beneath mid-ocean ridges, and more

enriched signatures in mantle that feeds hotspots, have led to the view that ancient crustal material has been recycled into mantle plume sources.

In particular, variations in lead (Pb) isotope ratios in oceanic basalts have been interpreted as evidence for recycling of oceanic crust into the mantle with timescales of around 2 billion years². For these heterogeneities to persist for billions of years despite convective stirring of the mantle, it is believed that the oceanic crust must undergo long-term storage at a boundary layer within the Earth, for instance at the core–mantle boundary, before ascending back through the convecting mantle³. But if some of the radiogenic Pb (produced by radioactive decay of uranium and thorium) now in mantle-plume sources was derived from ancient continental crust, then much of the 'ageing' of the Pb may have occurred before it entered the mantle. If so, the timescales of lithospheric recycling would be hundreds of millions rather than billions of years⁴.

The rhenium–osmium isotope system, based on the decay of ¹⁸⁷Re to ¹⁸⁷Os with a half-life of 42 billion years, provides a complementary view of lithospheric recycling. Because Re is moderately incompatible and Os strongly compatible during melting, mantle that has undergone extensive melt depletion is characterized by low Re/Os ratios, and with time develops low ¹⁸⁷Os/¹⁸⁸Os ratios compared with less melt-depleted mantle. A unique feature of the Re–Os system is its resistance to subsequent resetting, and its ability to 'remember' ancient melt-depletion events^{5,6}. So the strongly melt-depleted lithospheric mantle associated with basaltic oceanic crust should develop and retain distinctly low ('unradiogenic') ¹⁸⁷Os/¹⁸⁸Os signatures compared with

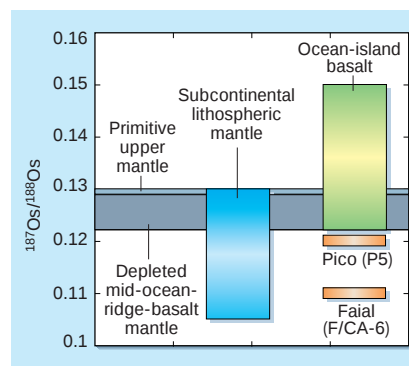


Figure 2 Osmium isotope signatures in different mantle reservoirs^{5,10,11}. Primitive upper mantle is mantle that has not experienced melt depletion. Depleted mid-ocean-ridge-basalt mantle is the shallow convecting source of mid-ocean-ridge basalts. Subcontinental lithospheric mantle is ancient, melt-depleted mantle from beneath continents. Ocean-island basalt is volcanic rock, found in places such as Hawaii and Iceland, formed by melting of mantle plumes. The isotope signatures of two ocean-island basalts from the Azores reported by Schaefer *et al.*¹ — Pico (P5) and Faial (F/CA-6) — are lower than any observed previously in oceanic rocks, from which the authors conclude that the source is ancient oceanic mantle lithosphere that was recycled and stored at depth. But these values are within the characteristic range evident in subcontinental lithospheric mantle.

the less melt-depleted convecting mantle.

This is where the results of Schaefer *et al.*¹ come in. They report Os isotope signatures in basalts from two islands in the Azores — Pico (Fig. 1) and Faial — that are lower than any previously found in oceanic basalts (Fig. 2). In short, these results indicate that the Azores plume contains a component of melt-depleted mantle that was stripped of its Re about 2.5 billion years ago. If this ancient mantle represents recycled oceanic lithosphere, as Schaefer *et al.* argue, then it must have undergone long-term storage in the mantle before being incorporated into the Azores plume — residence times of oceanic lithosphere on the Earth's surface are 200 million years at most, not billions of years.

In this case, either mantle mixing is much less efficient than most models suggest³, or the subducted slab must have been stored deep within the Earth for more than 2 billion years to develop the distinct unradiogenic Os isotope signatures inherited by the Azores plume basalts. Many lines of evidence support this second possibility in principle, and indicate that subducted oceanic lithosphere should penetrate the boundary between the upper and lower mantle at 660 km, and sink down to the core–mantle boundary. There it can remain isolated from the convecting mantle for long periods³.

Analysis of other isotope ratios in the Azores basalts may help to elucidate the



Figure 1 Pico Peak — the volcano on Pico Island in the Azores, the dramatic result of a mantle plume beneath the islands.

nature of the recycling process. The lower mantle (and possibly the core) has a greater complement of primordial ^3He , incorporated into the Earth at the time of accretion, because it has lost less of it through degassing than the shallow, upper-mantle source of mid-ocean-ridge basalts. High $^3\text{He}/^4\text{He}$ ratios relative to those in mid-ocean-ridge basalts are expected in basalts generated from mantle plumes that have risen through the lower mantle, because some entrainment of surrounding material will occur.

Significantly, then, basalts from Faial have $^3\text{He}/^4\text{He}$ ratios similar to and lower than those characteristic of mid-ocean-ridge basalts⁷. Low ratios would be expected for degassed, melt-depleted lithospheric mantle⁸ that has undergone only shallow recycling, and has not ascended through the lower mantle. So an alternative model to that proposed by Schaefer *et al.* is that the unradiogenic Os signatures in the Azores basalts are due to melt-depleted mantle from beneath ancient continents, which has peeled off and mixed with shallow convecting mantle⁹. In this case, long-term storage at the core–mantle boundary would not be required, because ‘ageing’ would have occurred at the base of stable continental lithosphere, which can be isolated from the convecting mantle for billions of years⁶.

A solution to this dilemma perhaps lies in

yet another isotope system, that of oxygen. If the Faial and Pico basalts were generated from recycled oceanic-mantle lithosphere, they would be expected to have low $^{18}\text{O}/^{16}\text{O}$ ratios characteristic of mantle that has experienced extensive high-temperature interactions with circulating sea water. Conversely, $^{18}\text{O}/^{16}\text{O}$ ratios similar to those in mid-ocean-ridge basalts would be expected if the source were lithospheric mantle derived from beneath continents, even if it had interacted with mantle fluids or melts⁹. Either outcome would provide additional insight into how the Earth’s interior evolves through time. ■

Elisabeth Widome is in the Department of Geology, Miami University, Oxford, Ohio 45056, USA.

e-mail: widome@muohio.edu

- Schaefer, B. F., Turner, S., Parkinson, I., Rogers, N. & Hawkesworth, C. *Nature* **420**, 304–307 (2002).
- Chase, C. G. *Earth Planet. Sci. Lett.* **52**, 277–284 (1981).
- van Keken, P. E., Hauri, E. H. & Ballentine, C. J. *Annu. Rev. Earth Planet. Sci.* **30**, 493–525 (2002).
- Hanan, B. B. & Graham, D. W. *Science* **272**, 991–995 (1996).
- Shirey, S. B. & Walker, R. J. *Annu. Rev. Earth Planet. Sci. Lett.* **26**, 423–500 (1998).
- Pearson, D. G., Carlson, R. W., Shirey, S. B., Boyd, F. R. & Nixon, P. H. *Earth Planet. Sci. Lett.* **134**, 341–357 (1995).
- Moreira, M., Doucelance, R., Kurz, M. D., Dupré, B. & Allègre, C. J. *Earth Planet. Sci. Lett.* **169**, 189–205 (1999).
- McKenzie, D. & O’Nions, R. K. *J. Petrol.* **36**, 133–159 (1995).
- Mattey, D., Lowry, D. & Macpherson, C. *Earth Planet. Sci. Lett.* **128**, 231–241 (1994).
- Meisel, T., Walker, R. J., Irving, A. J. & Lorand, J.-P. *Geochim. Cosmochim. Acta* **65**, 1311–1323 (2001).
- Walker, R. J., Prichard, H. M., Ishiwatari, A. & Pimentel, M. *Geochim. Cosmochim. Acta* **66**, 329–345 (2002).

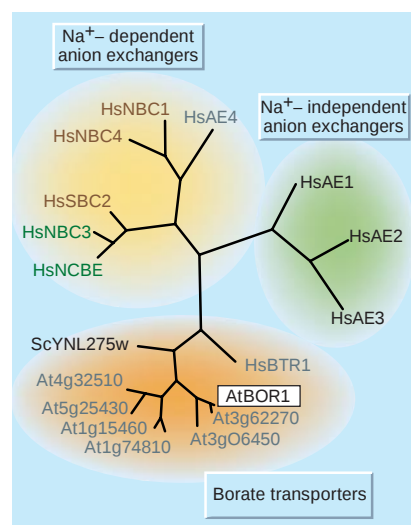


Figure 1 Phylogenetic analysis of the SLC4 anion-exchanger superfamily. The analysis was carried out by Daniel Wipf (University of Tübingen), using the program PAUP 4.0b10, and shows that the borate transporter identified by Takano *et al.*² falls in a clade with the yeast protein YNL275w, human BTR1 and six other *Arabidopsis* proteins. The other main clades are the mammalian Na^+ -independent and Na^+ -dependent anion exchangers. In the latter group, Na^+ -bicarbonate cotransporters are shown in brown, Na^+ -dependent anion exchangers in green. Proteins of unknown function are in grey. Hs, human; Sc, yeast; At, *Arabidopsis*.

Plant biology

Ping-pong with boron

Wolf B. Frommer and Nicolaus von Wirén

The identification of a transport mechanism for boron in plant roots provides a surprising connection with transport systems in other, very different settings, such as the kidney.

Boron is an essential nutrient for most organisms, and can be acquired from aqueous solution as boric acid. It permeates cell membranes easily, however, making it difficult to control¹, and plants and animals are subject to boron deficiency unless they counteract diffusion by active transport. Moreover, excess boron leads to toxicity, explaining its use in disinfectants and insecticides.

On page 337 of this issue, Takano *et al.*² report the first identification of a boron transporter, from *Arabidopsis thaliana*, a plant much favoured by experimental plant biologists. The transporter is responsible for exporting boron from specialized root cells into the xylem, the conduit in plants for carrying nutrients upwards from the roots.

Although the lipid membranes that enclose cells are designed to limit the entry and exit of solutes, many compounds are able to cross them. The rate at which that occurs is determined by several factors

— the properties and number of substrate-permeating transporters residing in the membrane, the properties of the substrate and the electrochemical potential difference across the membrane. Naturally, small hydrophobic molecules can pass through membranes more easily than larger or charged molecules. Not long ago, the prevailing opinion, mainly based on oil–water distribution coefficients, was that water, glycerol, ammonia, urea and weak acids such as carbonic, boric or acetic acid do not require transporter assistance. The surprise came when transporters such as aquaporins for water and urea, secondary active transporters for urea, and even fatty-acid transporters, turned out to be crucial whenever rapid flux is required. But it remained unclear whether boron uses a transporter. The answer is that it does.

The starting point for Takano *et al.*² was a mutant *Arabidopsis* plant that behaves as a *Japanese candy with two tastes* (hitotsubu-

de-nido-oishii). Previous work³ had identified this plant as a mutant in which virus replication is affected. But further research was hampered because the mutants did not set flowers and seeds, a typical characteristic of boron deficiency. Addition of borate crystals restored flowering and showed that the original line carried two independent mutations³ — one taste of the candy being a gene involved in defence against viruses; the other, a mutation in the *BOR1* gene, led to symptoms of boron deficiency.

Takano and colleagues' tracer studies revealed that roots of mutants contain adequate levels of boron, but that the plants suffer from reduced boron delivery to shoots because of impaired xylem loading. The mutation was mapped, and the *BOR1* gene was cloned and shown to be expressed in the pericycle, a ring of cells around the xylem that is responsible for exporting nutrients into xylem sap⁴. Takano *et al.* went on to provide a direct demonstration that *BOR1* transports boron, and that it is a membrane protein related to the family of mammalian anion exchangers known as SLC4 (Fig. 1). The similarity of these proteins and their substrates confers a much broader significance on the findings.

The prototype anion exchanger is BAND3, a transporter in red blood cells that is responsible for the exchange of

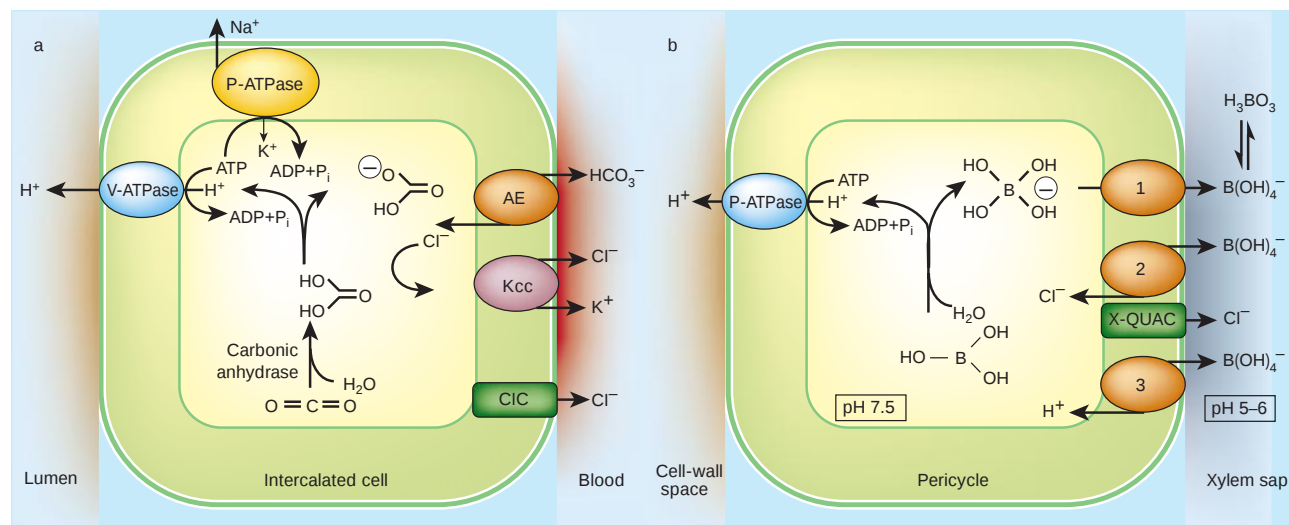


Figure 2 Comparative models of bicarbonate in the kidney and of borate in the xylem. **a**, In kidney, bicarbonate anions are formed from CO₂. Intercalated cells of the renal collecting tubule export this bicarbonate through an anion exchanger (AE) into the bloodstream, and a proton is exported into the lumen by a V-type ATPase⁵. Chloride is re-exported through K⁺/Cl⁻ cotransporters (Kcc) and CIC chloride channels^{6,7}, and potassium homeostasis and membrane potential are maintained by a P-ATPase. **b**, Borate export into the xylem from the root pericycle. In the

pericycle, boron loading of the xylem sap probably requires borate anions that are formed as a result of the high cytosolic pH. H⁺-ATPases acidify the cell-wall space and generate a negative membrane potential. The BOR1 transporter could export borate into the xylem in three different ways: (1) by diffusion of borate anions following the concentration gradient for borate (uniport); (2) by borate/chloride anion exchange coupled to a chloride gradient established by X-QUAC anion channels; or (3) by coupled countertransport (antiport) of borate with a proton.

bicarbonate and chloride using a ping-pong mechanism. In this context, ping-pong refers to a two-step process in which one anion binds on one side of the transporter pore, but exchange occurs only after the second substrate has bound on the other side. This mechanism is also used in pH regulation in the kidney⁵⁻⁷ (Fig. 2a). Other members of this family use a different mechanism in which they couple transport of sodium with that of bicarbonate in the same or the opposite direction⁵.

The surprising similarity of the transport systems for boron and bicarbonate points to a similarity in the binding forms of these two substrates. In plant cells a high cytoplasmic pH allows the formation of the borate anion, whereas in kidney cells bicarbonate formation from CO₂ is also enzymatically facilitated (Fig. 2). Given the phylogenetic relation between the proteins, the simplest hypothesis is that BOR1 also transports anions, and that perhaps borate transport is coupled to the antiport of a counterion in the same way as bicarbonate. BOR1 could then function as a borate/chloride anion exchanger using the chemical gradient established by certain chloride channels (X-QUAC channels)⁸. Alternatively, it could use proton coupling, instead of chloride coupling, to export borate by a secondary active route. Another possibility is that the negative membrane potential in the pericycle would allow borate anions to be exported by BOR1-mediated uniport (diffusion through a transporter without coupling to a second ion; Fig. 2b). Electrophysiological analyses of BOR1 in various settings should help to decide this matter.

The fully sequenced *Arabidopsis* genome has seven members in the anion-exchanger superfamily. They all fall into a single group (clade) along with the yeast protein YNL275w, the function of which is unknown. The two other clades in the superfamily are Na⁺-independent and Na⁺-dependent anion exchangers (Fig. 1). Phylogenetically, the yeast transporter seems to be an intermediate between the anion exchangers and BOR1, and so could potentially transport both bicarbonate and borate². Most importantly, the human BTR1 protein, named as being a possible bicarbonate-like transporter⁹, also falls into the same clade as BOR1 — it could keep the same abbreviated terminology if it in fact turns out to be a borate transporter.

The phylogenetic analysis gives no firm indication of the function of the six other *Arabidopsis* proteins shown in Fig. 1. However, some of them might serve as borate transporters that use other coupling mechanisms and that provide a route, for example, for importing boron into cells — one of the remaining questions. Many years ago, biophysical studies had indicated that bicarbonate and borate may use the same transporter¹⁰. Thus, some of the anion-exchange transporters similar to YNL275w may transport bicarbonate as well as borate, for example to facilitate the supply of CO₂ for photosynthesis. Study of this transporter family may therefore shed light not only on the functions of boron in metabolism but perhaps also on CO₂ movement in plants. Given the close relationship of bicarbonate and borate transporters among anion exchangers, it could be that relatives of the

active bicarbonate transporter from cyanobacteria may transport borate¹¹. Finally, some aquaporins may be permeable to borate and serve in boron transport¹².

The results of Takano *et al.*² may also have practical implications. The genetic differences in the susceptibility of plants to borate deficiency or toxicity¹³ have been attributed in part to variability in the lipid composition of their membranes¹⁴. But access to the transporter genes may enable gene-mapping and breeding programmes to produce boron-efficient crops, and the use of transgenic approaches may help to optimize plants for growth in soils with high levels of boron. ■

Wolf B. Frommer is at ZMBP, University of Tübingen, Auf der Morgenstelle 1, D-72076 Tübingen, Germany.
e-mail: frommer@zmbp.uni-tuebingen.de
Nicolaus von Wirén is at the Institute for Plant Nutrition, Hohenheim University, D-70593 Stuttgart, Germany.
e-mail: vonwieren@uni-hohenheim.de

1. Brown, P. H. *et al.* *Plant Biol.* **4**, 205–223 (2002).
2. Takano, J. *et al.* *Nature* **420**, 337–340 (2002).
3. Yamanaka, T. *et al.* *Proc. Natl Acad. Sci. USA* **97**, 10107–10112 (2000).
4. Gaymard, F. *et al.* *Cell* **94**, 647–655 (2002).
5. Alper, S. L., Darman, R. B., Chernova, M. N. & Dahl, N. K. *J. Nephrol.* **15**, S41–S53 (2002).
6. Boettger, T. *et al.* *Nature* **416**, 874–878 (2002).
7. Estevez, R. *et al.* *Nature* **414**, 558–561 (2001).
8. Köhler, B. *et al.* *Plant J.* **30**, 133–142 (2002).
9. Parker, M. D., Ourmozdi, E. P. & Tanner, M. J. *Biochem. Biophys. Res. Commun.* **282**, 1103–1109 (2001).
10. Lucas, W. J. *J. Exp. Bot.* **26**, 331–346 (1975).
11. Omata, T. *et al.* *Proc. Natl Acad. Sci. USA* **96**, 13571–13576 (1999).
12. Dordas, C. & Brown, P. H. *Biol. Trace Elem. Res.* **81**, 127–139 (2001).
13. Jeffries, S. P. *et al.* *Theor. Appl. Genet.* **101**, 767–777 (2000).
14. Dordas, C. & Brown, P. H. *J. Membr. Biol.* **175**, 95–105 (2000).

Obituary

Per Bak (1947–2002)

With Per Bak's death on 16 October 2002, Denmark lost one of its most prominent scientists. Per Bak was known worldwide for his contributions to the physics of complex systems and was one of the world's most cited physicists. His premature death at the age of 54 leaves a deep void at the Niels Bohr Institute, where his lively nature and rich stream of ideas never failed to engage the people around him.

Bak studied at the Technical University of Denmark and obtained his PhD from the Risø National Laboratory in 1974. Soon after, he moved to the Brookhaven National Laboratory in New York. His long association with Brookhaven was interrupted by stays at the University of Copenhagen and at the Nordic Institute of Theoretical Physics in Copenhagen. From the start, Bak showed himself to be a creative and productive scientist, concentrating in these early years on critical phenomena and phase transitions in metals and other materials. His 1976 work on the structure of synthetic one-dimensional conductors is a fine example of the elegance of his scientific methods and was for several years one of the most cited publications in solid-state physics.

In 1979 Bak was employed as a lecturer at the University of Copenhagen's Ørsted Laboratory and turned his attention to chaos and fractal theory. As usual, Bak was several steps ahead of everyone else, and his ability to combine different areas of physics enabled him to lead this developing field into uncharted territory. His skill at generating original scientific ideas gave him the opportunity to travel to conferences and universities around the world. Although he had a great talent for guiding students, he was not totally committed to curricular teaching. Rather, he would prefer to find thought-provoking projects for his students and spent days discussing them. In 1982, Bak received a doctor of science degree for his thesis on modulated and chaotic phases — another classic, highly cited paper.

But sometimes he felt that Denmark was too small to satisfy his restless nature, and in 1983 Bak returned to a permanent position at Brookhaven. This was the start of another productive period, in which he turned his attention to what is now called the physics of complex systems. In 1987, with Chao Tang and Kurt Wiesenfeld, he published what he and many colleagues considered to be his most important contribution to science: the theory of self-organized criticality (SOC). For years Bak



Pioneer in the physics of complex systems, and discoverer of self-organized criticality

had pondered the many examples of fractal structures found in nature — from clouds and earthquakes to the distribution of galaxies in the Universe — and his theory of SOC underpins the observed self-similar structure of such systems. Once again, he had demonstrated his particular ability to piece together observations and theoretical knowledge to make a great leap forward.

The theory of self-organized criticality spread, like the forest fires it describes, through the scientific community, and his ideas were adopted around the world. As with any of Bak's proposed theories, there were sceptics as well as adherents, but the latter were far more numerous. Bak enjoyed the often heated debate and participated enthusiastically in discussions at conferences and by e-mail. He never went with the flow but threw himself with tremendous energy against it. For those around him, this attitude brought opportunities to participate in exciting discussions, with fresh ideas always seeming to ferment just below the surface.

Bak's lively temper was legendary. Woe betide that seminar speaker who had misunderstood a central idea or happened

to find himself at cross-purposes to it. Bak's objections were always delivered with a touch of humour and were mostly well founded.

Over the past 15 years, the theory of SOC has become an established part of many scientific disciplines, and remains one of the most cited works in theoretical physics. Bak was delighted that his theory could be applied to virtually every branch of science and that it stands as one of the most successful interdisciplinary contributions to science in recent decades. He collaborated and published with economists, geologists, biologists, physicians, mathematicians and many others. The theory of SOC was also the theme of his popular book *How Nature Works* (Copernicus, 1996), which has been translated into several languages.

Bak was a frequent guest speaker at international conferences on statistical physics and the physics of complex systems. His lecturing style was always relaxed, even when he was putting forward the often controversial ideas from his latest research. He was interested in the big picture. Although details did not concern him, they were often the source of questions from his audience, and his elegant brush-offs were a source of great amusement for his listeners. A lecture by Bak was never boring, and conference organizers all over the world knew it.

In 1996, Bak was called home to take up a professorship in theoretical physics at the Niels Bohr Institute in Copenhagen. He greatly appreciated this recognition, and valued the unique atmosphere at the institute. But he watched with growing concern as the institute was reduced by innumerable budget cuts. He viewed the hierarchical management and control mechanisms of politicians and bureaucrats with open contempt, and was convinced that heavy-handed management of science would strangle free-thinking research, replacing it with 'programmes' of little originality and no lasting value. The rare but genuinely revolutionary advances would succumb, he feared, to a steady stream of uninteresting technical progress.

Over the past few years, Bak believed that Denmark's reputation in fundamental research was in serious decline, in spite of his own considerable efforts to further it. Indeed, the recognition of truly free scientific research as a central element in any healthy society would be the most fitting memorial to Per Bak's own life and work.

Mogens Høgh Jensen

Mogens Høgh Jensen is at the Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, 2100 Copenhagen, Denmark.
e-mail: mhjensen@nbi.dk

Archaeoraptor's better half

The other component of this infamous fossil forgery is identified as a fish-eating bird.

The 'Archaeoraptor' fossil, once proclaimed as a key intermediate between carnivorous dinosaurs and birds¹ but now known to be a forgery, is a chimera formed of bird and dromaeosaur parts^{2,3}. Although the non-avian dinosaur tail of this controversial specimen from the Early Cretaceous Period of China has been identified⁴, the avialan parts of the specimen have not. Here we reveal that these avialan parts, including the hindlimbs, which were previously designated as unverifiable "attached bones"³, can be referred to a single species, *Yanornis martini*⁵, and are supported as pieces of a single, almost complete specimen. A new specimen that is also referable to this species, and which has its gut contents preserved, indicates that the principal part of this false raptor dinosaur–bird fossil is in fact a fish-eating bird.

High-resolution computed X-ray tomography revealed that two to five specimens belonging to two or more species could comprise the *Archaeoraptor* forgery³. However, the morphology and proportions of all of these parts, other than the dromaeosaur tail, supports referral to a single species, *Y. martini*⁵. Measurements (see supplementary information) and anatomical features of the *Archaeoraptor* parts correspond closely and, taking breakage into account, are almost identical to those of the *Yanornis* holotype.

The premaxillae are partially fused and toothed in both specimens (Fig. 1). The tip of the exposed right premaxilla of the *Archaeoraptor* specimen lacks associated teeth. The length of this portion of the facial margin is identical to that of the edentulous portion of the premaxilla in the *Yanornis* holotype. In both specimens, the coracoid has a prominent procoracoid process and deeply concave scapular cotyla (Fig. 1). These skull and pectoral-girdle features are preserved as impressions in the *Archaeoraptor* specimen (Fig. 1a). Both preserve prominent lateral processes of the coracoids, as well as extremely robust furculae ('wishbones'). The combination of the aforementioned morphologies is unique to *Yanornis*, and the dimensions of these elements are equivalent (Fig. 1b).

The dimensions of the femur, tibia and tarsometatarsus, as well as the lengths of the pedal phalanges, are also nearly identical and are consistent with referral of these parts to *Yanornis* (see supplementary information). The foot shares with the *Yanornis* holotype morphological features, including proximal and distal fusion of metatarsals II–IV, and metatarsal II being shorter than IV.

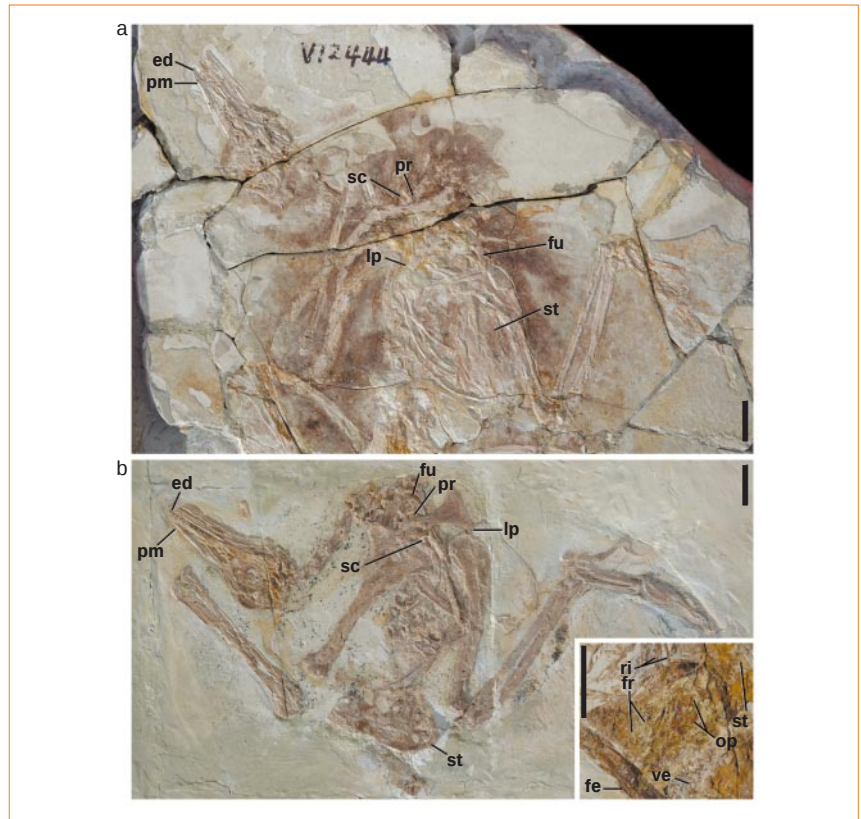


Figure 1 Analysis of the avialan half of the 'Archaeoraptor' forgery. **a**, Front half of the *Archaeoraptor* specimen IVPP V12444 (excluding the tail); **b**, front half of the *Yanornis martini* holotype specimen IVPP V13259. ed, edentulous portion of premaxilla; fe, femur; fr, fin rays; fu, furcula; lp, lateral process of coracoid; op, teleost fish opercular fragments; pm, premaxilla; pr, procoracoid process of coracoid; ri, ribs; sc, scapular cotyla of coracoid; st, sternum; ve, teleost fish vertebra. Scale bars, 2 cm. For further data used in the analysis, see supplementary information.

On the basis of the presence of a combination of morphologies unique to *Y. martini* in the associated front part of the *Archaeoraptor* specimen listed here, we refer this portion of the specimen to that taxon. Because of the correspondence in measurements and morphology, the hindlimb elements are identified as being consistent with these separate blocks that make up parts of a single *Yanornis* specimen. The avialan portion of the *Archaeoraptor* specimen also provides further evidence of the anatomy of *Y. martini*. For example, it preserves a distally tapering scapular blade, a short, slightly recurved acromion, and remnants of an anteriorly developed sternal keel, features that are not preserved in the holotype.

A recently discovered specimen of *Yanornis* (IVPP V13259) contains preserved macerated fish remains, including a teleost vertebra, fin rays and opercular fragments (Fig. 1b, inset), and is only the second bird from the exceptionally diverse Jehol Biota^{6,7}

to be discovered with its gut contents preserved. The principal portion of the *Archaeoraptor* forgery, which we conclude was constructed from two different specimens belonging to two different species, is therefore representative of a fish-eating bird.

Zhonghe Zhou*, Julia A. Clarke†, Fucheng Zhang*

**Institute of Vertebrate Paleontology and Paleoanthropology, PO Box 643, Beijing 100044, China
e-mail: zhonghe@yeah.net*

†Section of Vertebrate Paleontology, American Museum of Natural History, New York, New York 10024, USA

1. Sloan, C. P. *Natl Geogr.* **196**, 98–107 (1999).
 2. Dalton, R. *Nature* **403**, 689–690 (2000).
 3. Rowe, T. et al. *Nature* **410**, 539–540 (2001).
 4. Xu, X., Zhou, Z. & Wang, X. *Nature* **408**, 705–708 (2000).
 5. Zhou, Z. & Zhang, F. *Chinese Sci. Bull.* **46**, 371–377 (2001).
 6. Zhang, M., Chen, P., Wang, Y.-Q. & Wang, Y. (eds) *The Jehol Biota* (Shanghai Sci. Technol., Shanghai, 2001).
 7. Zhou, Z. & Zhang, F. *Nature* **418**, 405–409 (2002).
- Supplementary information** accompanies this communication on Nature's website.
Competing financial interests: declared none.

DNA nanotechnology

Chemical copying of connectivity

Three-dimensional DNA nanoscaffolds such as supramolecular tetrahedra can self-assemble from tris-oligonucleotidyls — synthetic three-armed building blocks in which three identical or non-identical short DNA sequences are connected by a tris-linking backbone^{1,2}. Here we show that the connectivity information contained in these building blocks can be copied by using template-directed tris-linking. This finding is a crucial step towards the replication of nanoarchitectures that are based on tris-oligonucleotidyls and to the realization of artificially self-replicating systems on a nanometre scale.

The procedure is outlined in Fig. 1a. On the basis of the three template sequences of an asymmetric tris-15-nucleotide oligomer (ref. 2), we generated the three corresponding complementary linear sequences with a reactive group at their 5' ends, and used these in different ligation reactions^{3–5} with a suitable tris-linker molecule.

Template-directed tris-linking was rapid and robust — unhampered by hydrolysis — in the hydrazone-formation reaction shown in Fig. 1b. Three 5'-hydrazide-modified oligonucleotides, a 15-nucleotide molecule (15-mer; **A**), a 13-mer (**B**) and an 11-mer (**C**), the sequences of which are complementary to the respective arms of the tris-15-meric template **Y**, were synthesized by using standard solid-phase protocols. The 5'-hydrazide moieties were introduced by using a modifier amidite⁶.

For a statistical, non-template-directed synthesis, ten 5'-tris-linked species would be expected, ranging from a branched 33-mer to a branched 45-mer, as well as six bis-linked species containing between 22 and 30 nucleotides. A template-directed synthesis should yield a single branched 39-mer, as well as three of the six possible bis-linked species, if the reaction has not reached completion. Figure 1c shows that the predicted 39-mer is formed as the principal product in the presence of the template (confirmed by MALDI mass spectroscopy); in the absence of the template, there are no tris-linked products detectable by polyacrylamide-gel electrophoretic analysis. There is also evidence that the template directs the formation of the intermediate bis-linked species, the distribution of which differs from that in the statistical synthesis.

The copies are stable, even without reductive amination in the presence of sodium cyanoborohydride, which has been used successfully in other imine-based, template-directed ligation reactions⁷. We also found that short-armed tris-9-mers

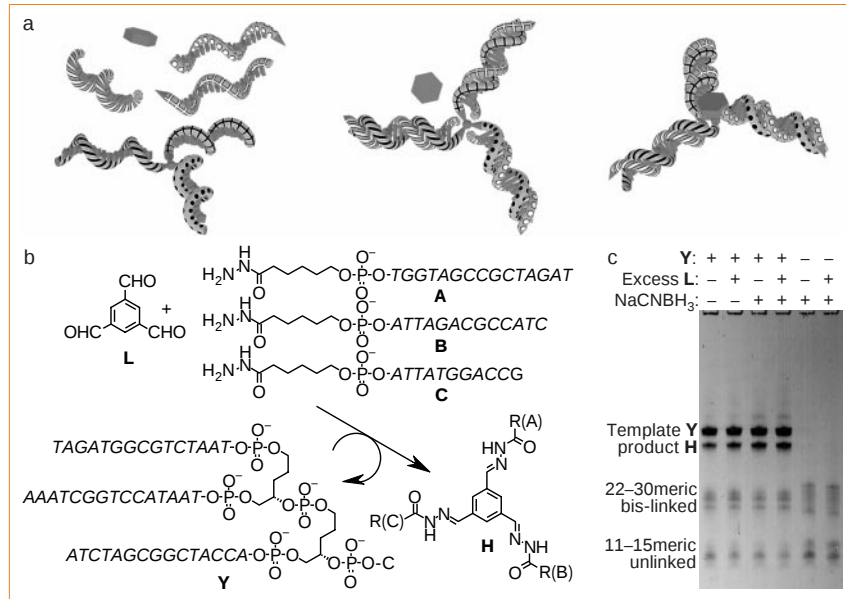


Figure 1 Chemical copying of connectivity by template-directed tris-linking. **a**, Left, chemical copying of connectivity involves a 3'-tris-oligonucleotidyl template with three individually defined sequences, three linear complements with 5' ends represented as blunt ends, and a tris-linker shown as a honeycomb cap. Centre, hybridization gives rise to a quartermolecular complex in which the blunt ends come into close spatial proximity. Right, the tris-linker has reacted to connect the 5' ends in the copy. **b**, Chemical copying of connectivity on the basis of tris-hydrazone formation. Tris-linking of the 5'-hydrazide-modified oligonucleotides **A**, **B** and **C** with the tris-linker **L** in the presence of the 3'-tris-oligonucleotidyl template **Y** yields the tris-hydrazone **H**. **c**, Nucleic-acid staining (with SYBR gold), after denaturing polyacrylamide-gel electrophoresis, of an almost complete chemical copying of connectivity from a 10 μ M solution of the components in 100 mM MOPS buffer (pH 5.6) after incubation for 2 h at room temperature in the presence or absence of template **Y**, an excess (2.5:1.0 equivalent) of tris-linker **L**, and sodium cyanoborohydride (NaCNBH₃; 100 equivalents).

can be used as templates to generate asymmetrical tris-30-mers, which are otherwise not easily accessible by chemical synthesis. Tris-linking was also possible with 3+1-armed templates², in which one arm is connected through its 5' end. The 'Klenow' polymerase enzyme accepts the 3' end of tris-15-mer copies hybridized to overhanging linear templates as extension sites.

A full replication cycle also requires reverse copying — that is, the tris-linking of 3'-modified oligonucleotides using 5'-tris-oligonucleotide templates. Once this step becomes feasible, surface-promoted replication and exponential amplification of DNA analogues (SPREAD)⁸ may be adapted to connectivity templating to achieve multiple rounds of replication. For self-assembling, non-covalent nanostructures, including tetrahedral nanoscaffolds of modular functions², replication will depend only on the ability of the subunits to replicate. Non-covalent connectivity therefore constitutes transferable information that can instruct a subunit where to integrate during self-assembly. Covalent connectivity and/or connectivity by topological interweaving, as used in Seeman's DNA cube⁹, cannot replicate

but ensures ultimate stability. Switchable duplex crosslinkers, such as those based on disulphide bonds¹⁰, may be suitable for stabilizing replicatable nano-objects.

Lars Henning Eckardt*, Kai Naumann*, Wolf Matthias Pankau*, Michael Rein*, Markus Schweitzer†, Norbert Windhab†, Günter von Kiedrowski*

*Lehrstuhl für Bioorganische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany
e-mail: kiedro@ernie.orch.ruhr-uni-bochum.de
†Nanogen Recogomics, Industriepark Hoechst, 65926 Frankfurt, Germany

- Scheffler, M., Dorenbeck, A., Jordan, S., Wüstefeld, M. & von Kiedrowski, G. *Angew. Chem. Int. Ed.* **38**, 3311–3315 (1999).
- Dorenbeck, A., Scheffler, M., Wüstefeld, M. & von Kiedrowski, G. *Angew. Chem.* (in the press).
- Dawson, P. E., Churchill, M. J., Ghadiri, M. R. & Kent, S. B. H. *J. Am. Chem. Soc.* **119**, 4325–4329 (1997).
- Stetsenko, D. A. & Gait, M. J. *J. Org. Chem.* **65**, 4900–4908 (2000).
- Cochran, J. R. & Stern, L. J. *Chem. Biol.* **7**, 683–696 (2000).
- Raddatz, S. et al. *Nucleic Acids Res.* **30**, 4793–4802 (2002).
- Li, X., Zhan, Z. Y. J., Knipe, R. & Lynn, D. G. *J. Am. Chem. Soc.* **124**, 746–747 (2002).
- Luther, A., Brandsch, R. & von Kiedrowski, G. *Nature* **396**, 245–248 (1998).
- Chen, J. & Seeman, N. C. *Nature* **350**, 631–633 (1991).
- van Aalten, D. M. F., Erlanson, D. A., Verdine, G. L. & Leemor, J. T. *Proc. Natl Acad. Sci. USA* **96**, 11809–11814 (1999).

Competing financial interests: declared none.

brief communications is intended to provide a forum for both brief, topical reports of general scientific interest and technical discussion of recently published material of particular interest to non-specialist readers. Priority will be given to contributions that have fewer than 500 words, 10 references and only one figure. Detailed guidelines are available on *Nature's* website (www.nature.com/nature) or on request from nature@nature.com

Insights into DNA recombination from the structure of a RAD51–BRCA2 complex

Luca Pellegrini*, David S. Yu†‡, Thomas Lo*‡, Shubha Anand†, MiYoung Lee†, Tom L. Blundell* & Ashok R. Venkitaraman†

* University of Cambridge, Department of Biochemistry, Tennis Court Road, Cambridge CB2 1GA, UK

† University of Cambridge, CR UK Department of Oncology and The Medical Research Council Cancer Cell Unit, Hutchison/MRC Research Centre, Hills Road, Cambridge CB2 2XZ, UK

‡ These authors contributed equally to this work

The breast cancer susceptibility protein BRCA2 controls the function of RAD51, a recombinase enzyme, in pathways for DNA repair by homologous recombination. We report here the structure of a complex between an evolutionarily conserved sequence in BRCA2 (the BRC repeat) and the RecA-homology domain of RAD51. The BRC repeat mimics a motif in RAD51 that serves as an interface for oligomerization between individual RAD51 monomers, thus enabling BRCA2 to control the assembly of the RAD51 nucleoprotein filament, which is essential for strand-pairing reactions during DNA recombination. The RAD51 oligomerization motif is highly conserved among RecA-like recombinases, highlighting a common evolutionary origin for the mechanism of nucleoprotein filament formation, mirrored in the BRC repeat. Cancer-associated mutations that affect the BRC repeat disrupt its predicted interaction with RAD51, yielding structural insight into mechanisms for cancer susceptibility.

Germline mutations in the *BRCA2* gene cause increased susceptibility to breast, ovarian and other cancer types¹. The 3,418-amino-acid BRCA2 protein localizes to the nucleus during S phase of the mitotic cell cycle, and is also highly expressed during meiosis. BRCA2 does not closely resemble other proteins of known function, yielding few clues to its biological function.

BRCA2 contains² eight conserved sequence motifs—the BRC repeats—of about 30 amino acids each, positioned between residues 990 and 2100 in human BRCA2. Although the overall sequence similarity among BRCA2 orthologues is limited, the BRC repeats in different species are well conserved^{3,4}, suggesting that they are essential to the biological functions of BRCA2. Indeed, the BRC repeats are the primary sites through which BRCA2 binds to RAD51 (refs 5–7), a protein with a crucial role in DNA recombination. Similar to the bacterial homologue RecA, RAD51 coats single-stranded DNA substrates to form a helical nucleoprotein filament, which can invade duplex DNA and pair with homologous nucleotides to initiate the strand exchange reactions that culminate in genetic recombination. When expressed *in vitro*^{5–7}, six of the eight BRC repeats in BRCA2 can interact directly with recombinant RAD51. BRC3 and BRC4 encoded in human BRCA2 are particularly efficient at RAD51 binding, whereas BRC5 and BRC6 are not.

The interaction between BRCA2 and RAD51 is critical for the biological functions of both molecules^{8,9}. Discrete nuclear foci containing RAD51 usually accumulate within the nucleus of mammalian cells exposed to DNA damage. RAD51 foci fail to form in BRCA2-deficient cells^{7,10,11}, suggesting that BRCA2 transports RAD51 to sites where DNA damage is processed by recombination. Indeed, BRCA2 deficiency leads to a severe defect in the repair of DNA double-strand breaks by recombination¹², and, similar to RAD51 deficiency^{13,14}, provokes spontaneous instability of chromosome structure during cell division^{15,16}. Notably, and in apparent conflict with these data, the activity of RAD51 in nucleoprotein filament formation is suppressed by its interaction with peptides encoding BRC repeats¹⁷. Collectively, the experimental evidence suggests models in which the cellular localization of BRCA2–RAD51 complexes and their activity in nucleoprotein filament formation are regulated after DNA damage, perhaps resulting in

transitions from ‘inactive’ to ‘active’ states^{9,17}.

The functional complexity of the interaction between BRCA2 and RAD51 prompted us to take a structure-based approach to analysis. Here, we report the 1.7 Å crystal structure of a complex between BRC repeat 4 (BRC4) and the RecA-homology domain of RAD51. To overcome the natural tendency of RAD51 to form heterogeneous aggregates, we covalently linked BRC4 to RAD51 to enable purification of the complex as a fusion protein. Our data reveal the structural basis for the BRCA2-dependent regulation of RAD51 function in DNA recombination, and provide structural insight into BRCA2 mutations associated with increased susceptibility to cancer.

Architecture and interface of the RAD51–BRC4 complex

Crystallographic data are summarized in the Supplementary Information. The structure of the human RAD51 ATPase domain, the first to be reported for a RAD51 protein, confirms that RAD51 belongs to the RecA-like family of ATPases (Fig. 1). The catalytic domains of human RAD51 and bacterial RecA¹⁸ are topologically identical and their superposition results in a root mean square deviation of 1.7 Å over 160 Cα atoms (out of 210 present in the crystallographic model).

BRC4 remains in continuous contact with RAD51 over a stretch of 28 amino acids (Leu 1521 to Glu 1548) (Fig. 1a, b). Residues Phe 1524 to Val 1532 form a β-hairpin with a 3:5 loop (1526–TASGK-1530) structured as a type I turn followed by a β-bulge at residue Gly 1529, which has a positive ϕ torsion angle¹⁹. The hairpin lines up alongside β-strand B3, extending the β-sheet of RAD51 by two short anti-parallel strands. After the hairpin, the BRC motif wraps around helix A4 of RAD51 through a short linker (Lys 1533 to Ala 1535) that leads into an amphipathic α-helical segment (Lys 1536 to Val 1542). The remaining residues at the carboxyl end of BRC4 (Lys 1543 to Glu 1548) form an irregular coil with elements of a 3₁₀ helix that spans helices A4 and A5 of RAD51, making an angle of 60° to their axes.

The RAD51–BRC4 complex buries 2,026 Å² of surface area, with substantial hydrophobic interactions, but also polar contacts, present at the interface (Fig. 2). The BRC motif is decorated with hydrophobic residues that keep it in close contact with RAD51

throughout its length (Fig. 2a, b). Three main points of contact stand out, involving residues Phe 1524, Ala 1527, Leu 1545 and Phe 1546. Phe 1524 is located on the strand of the β -hairpin contacting RAD51, and its aromatic ring is buried within a hydrophobic cavity formed by the side chains of RAD51 residues Met 158, Ile 160, Ala 190, Ala 192, Leu 203, Ala 207 and Met 210. Ala 1527, in the second position of the hairpin loop, places its β -carbon into a small pocket formed by the side chains of RAD51 residues Phe 166, Pro 168, Leu 171, Leu 186 and Val 189. In the C terminus of the BRC repeat, Leu 1545 and Phe 1546 form a wedge embedded between RAD51 helices A4 and A5, and surrounded by residues Leu 204, Tyr 205, Ser 208 (in helix A4), and Met 251, Arg 254, Leu 255, Glu 258 and Phe259 (in helix A5). The affinity between BRC4 and RAD51 is further enhanced by hydrophobic contacts involving residues Ile 1534 in the linker region, and the hydrogen-bonded Ser 1538, Leu 1539 and Val 1542 in the α -helix.

Although not as numerous as the hydrophobic interactions, contacts of a polar and charged nature also occur (Fig. 2c). The β -hairpin keeps BRC4 in register relative to RAD51 through a set of three continuous, anti-parallel main-chain to main-chain hydrogen bonds linking the BRC4 sequence 1525-HTA-1527 to the 189-VAY-191 sequence in strand B3 of RAD51. Asp 187 of RAD51 accepts a hydrogen bond from Ser 1528, the third position of the BRC4 hairpin loop, and interacts electrostatically with Lys 1530. Glu 213 of RAD51 receives a hydrogen bond from Ser 1538 of BRC4, in what probably represents a particularly significant contact, as the two side chains are kept poised for interaction. The position of the Ser 1538 side chain is determined by a stacking interaction with BRC4 Ala 1535 and RAD51 Val 212, whereas Glu 213 is hydrogen-bonded to the main-chain nitrogen of Ala 1535 and, via a water molecule, to the main-chain carbonyl of Lys 1533. Finally, Glu 1548, at the C-terminal end of the BRC4 motif, forms an ion pair with Arg 250 of RAD51.

Interactions involving residues that are not conserved across BRC repeats help to explain the higher affinity⁷ of BRC4 towards RAD51 relative to other repeat types. Leucines 1521 and 1522 are in hydrophobic contact with the side chains of RAD51 residues Phe 195 and His 199, and the main-chain carbonyl of Leu 1522 accepts a hydrogen bond from His 199. His 1525 forms a pseudo-hydrophobic core by packing against the aliphatic portions of Lys 1531 and Thr 1520 side chains, thus conferring additional stability to the β -hairpin structure.

The conformation of the nucleotide-binding site (P-loop) differs between RAD51 and RecA (Supplementary Information). A three-dimensional superposition shows that, whereas the P-loop remains unchanged in apo- and ADP-bound RecA^{18,20}, in BRCA2-bound RAD51 it adopts a closer conformation that probably precludes its occupation by ATP. Although the BRC repeat does not mask directly the nucleotide-binding site, it may inhibit ATP binding through an indirect conformational effect. The physiological relevance of this observation is currently unclear.

BRC sequence conservation and cancer-associated mutations

The structure of the RAD51–BRCA2 BRC4 complex permits the rationalization of the pattern of sequence conservation observed in BRC repeats belonging to different repeat types and organisms (Table 1). Furthermore, it explains the potentially carcinogenic nature of point mutations affecting BRCA2 residues important for RAD51 binding (Breast Cancer Information Core database; <http://research.nhgri.nih.gov/bic/>).

The most amino-terminal residue to be significantly conserved, Gly 1523, resides in a spatially constrained environment at the protein–protein interface, and glycine or serine account for 60% of occurrences at this position. Residues 1524–FHTASGK-1530, with the exception of His 1525, form a contiguous block of highly conserved amino acids. The structure shows that they are involved

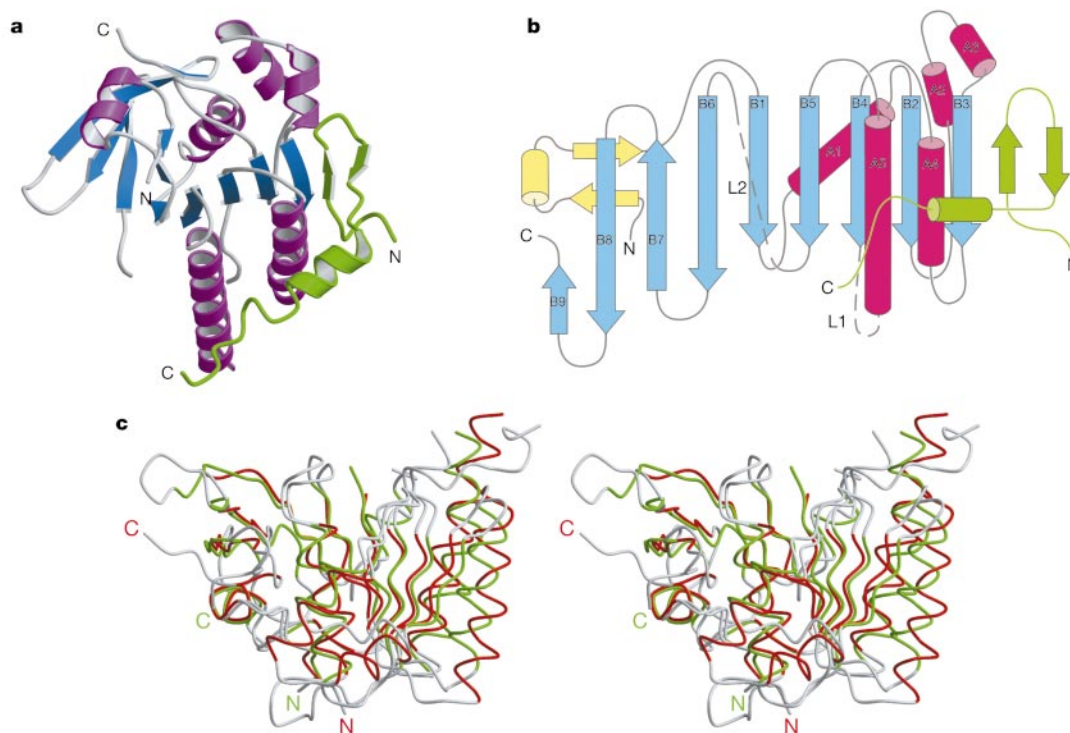


Figure 1 The RAD51–BRC4 complex. **a**, Ribbon representation. RAD51 is in magenta (α -helices) and light blue (β -strands), the BRC repeat motif is in green. The N and C termini are indicated. **b**, Topology diagram. The colour scheme is the same as in **a**, except for the N-terminal strand-helix-strand motif of RAD51 (yellow). Labelled secondary structure elements in RAD51 constitute the RecA-homology domain. Disordered RAD51

loops (dashed lines) connect β -strand B4 to α -helix A5 (L1) and B5 to B6 (L2). **c**, Stereo diagram superposing the ATPase domains of human RAD51 and bacterial RecA (Protein Data Bank accession code 2reb). Secondary structure elements are coloured red (RAD51) or green (RecA).

in critical hydrophobic (Phe 1524 and Ala 1527, present in 89% and 82% of BRC sequences, respectively) and polar (Ser 1528 and Lys 1530, 59% and 70%, respectively) interactions with RAD51.

A Thr1526Ala point mutation impairs the ability of a BRC4

peptide to bind RAD51 (ref. 7). Cancer-associated mutations affecting the equivalent position in BRC1 (T1011R), BRC2 (S1221P) and BRC7 (T1980I) have been reported. The structure shows that Thr 1526 accepts a hydrogen bond from the main-chain

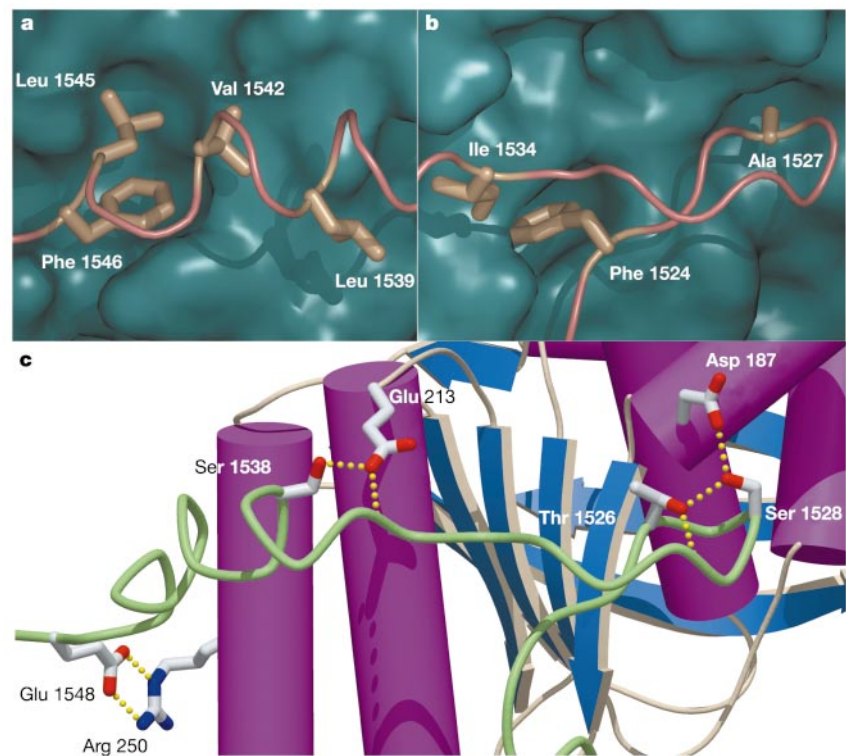


Figure 2 The RAD51–BRC4 interface. **a**, Hydrophobic interactions mediated by the α -helical region of the BRC repeat. RAD51 is rendered as a solvent-accessible molecular surface (green). The BRC motif is shown as a maroon coil, featuring only the relevant side chains (light brown). **b**, Hydrophobic interactions involving the β -hairpin region of the BRC

repeat. **c**, Polar interactions. RAD51 is shown in ribbon representation (colouring as in Fig. 1a), and the BRC motif is shown as a green tube. Interacting side chains are drawn in stick representation. Dotted yellow lines represent hydrogen bonds.

Table 1 Structure-based analysis of BRCA2 BRC repeat sequence conservation

Amino acid	BRC4 sequence																							
	β						β						α						α					
	L	L	G	F	H	T	A	S	G	K	K	V	K	I	A	K	E	S	L	D	K	V	K	N
D	2	-	-	-	1	-	2	3	-	1	-	-	-	-	-	8	1	-	-	6	2	2	-	1
E	1	1	-	-	-	-	1	-	1	9	-	2	-	-	-	21	22	-	-	3	4	1	1	4
K	1	-	1	-	1	-	-	1	-	39	20	-	12	1	-	10	2	-	1	8	44	-	27	12
R	-	4	-	1	8	-	-	1	1	5	2	-	4	2	1	2	-	-	4	1	-	11	1	1
H	-	-	-	-	6	-	-	6	1	-	-	-	-	-	-	3	4	-	-	5	-	1	-	-
N	3	1	1	-	-	-	-	1	7	-	7	1	4	-	-	2	1	6	-	5	1	2	1	11
Q	-	1	-	1	8	1	-	-	-	3	5	-	7	-	-	5	1	-	-	8	-	-	-	8
S	3	7	12	-	17	7	2	33	8	-	3	-	13	-	39	2	13	28	-	4	2	11	-	3
T	-	2	2	-	2	45	1	1	1	5	-	8	5	-	5	2	3	3	-	3	2	-	-	1
G	5	15	21	-	-	-	2	4	37	1	-	-	-	-	2	-	1	-	-	-	3	1	6	-
A	6	-	7	-	-	1	46	-	-	-	5	-	-	-	4	-	8	15	-	3	-	19	1	1
P	5	4	-	-	-	-	1	-	-	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-
C	3	-	-	2	3	-	-	2	-	-	-	1	5	-	-	-	-	-	1	-	-	2	-	-
I	-	4	5	-	3	1	-	-	-	-	2	16	-	-	-	-	2	8	-	-	3	-	-	6
L	8	16	-	-	-	1	-	3	-	-	2	6	-	7	1	-	1	38	3	-	-	1	-	17
V	2	1	4	-	-	-	1	1	-	1	-	23	2	33	-	1	-	-	4	1	-	14	7	-
M	-	-	-	-	-	-	-	-	-	-	-	-	-	4	-	-	-	-	-	-	4	7	6	-
F	14	-	3	50	-	-	-	-	-	-	-	1	1	1	-	-	-	1	4	2	-	-	1	8
W	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6	-
Y	3	-	-	2	7	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	1	-
Consensus sequence			G	F	x	T	A	S		G	S	K	o	i	x	i	S	o	o	S	L	x	K	A
			S			S				N										A			V	R
																								S

The BRC4 sequence from Leu 1521 to Glu 1548 is displayed horizontally across the top of the table. Residues within elements of secondary structure are boxed: α and β above the boxes indicate β -strand and α -helix, respectively. The twenty different amino acids are shown in the left column, grouped according to their chemical nature (hydrophilic at the top, hydrophobic at the bottom, the rest in the middle). Each value indicates the number of times a certain type of amino acid occurs at a particular position in the BRC repeat. The table contains sequence information relative to a set of 56 BRC repeats from 7 different organisms. The information contained in the table is recapitulated by the BRC consensus sequence reported under it. i, hydrophobic; o, hydrophilic; a, aromatic; x, no preference.

nitrogen of Lys 1530 that is essential for the conformational stability of the hairpin loop. Thus, BRC repeats in which the threonine is mutated are unlikely to assume the correct hairpin conformation necessary for efficient binding to RAD51. Thr 1526 also donates a hydrogen bond to the hydroxyl function of Ser 1528, thus keeping it poised for interaction with Asp 187 of RAD51. Threonine or serine account for 93% of occurrences at position 1526.

A common breast-cancer-associated mutation changes Gly 1529, at the fourth position of the hairpin loop, to arginine. Conformational restraints on position 1529 lead to selection of amino acids able to adopt a positive ϕ torsion angle, and glycine, serine or asparagine are found here with a combined frequency of 93%. Replacement of glycine by arginine will probably disrupt the conformation of the BRC β -hairpin, thereby weakening RAD51 binding.

In the linker connecting the β -hairpin to the α -helix, Val 1532 and Ile 1534 contribute to the continuous adherence of the BRC4 motif to the RAD51 surface through a hydrophobic contact with Met 210 of RAD51 (80% and 93% are the respective frequencies in these two positions for isoleucine, leucine or valine). A serine is found with 70% occurrence at position 1535, which caps the α -helix at its N terminus. Within the amphipathic helix, conserved residues including Ser 1538 (50% preference) and Leu 1539 (89% combined preference for leucine, isoleucine or valine) make hydrophobic and hydrogen-bonded interactions with RAD51. The selection of a small amino acid capable of hydrophobic interaction at position 1542 (79% combined frequency for valine, alanine or serine) is explained

by the close contact between BRC4 and helix A4 of RAD51. The strong preference for lysine at position 1541 (79%) is intriguing as its side chain is solvent-exposed, and suggests that BRC sequence conservation is not exclusively dictated by interaction with RAD51.

Leu 1545 and Phe 1546 are part of a large hydrophobic interface to which residues on helices A4 and A5 of RAD51 also contribute, and indeed hydrophobic residues are strongly represented at these positions (89% and 93% conservation, respectively). The structure further demonstrates that, whereas Leu 1545 is partially solvent exposed, the side chain of Phe 1546 penetrates deep into the RAD51–BRC4 interface. Accordingly, position 1545 shows only a general hydrophobic preference, whereas position 1546 requires either a phenylalanine or a leucine. The most C-terminal position to show a distinct sequence preference is 1548, which selects for an acidic residue (80% combined conservation for aspartic or glutamic acid). In the crystal, Glu 1548 forms a salt link with Arg 250 of RAD51.

The BRC motif is reminiscent of a Velcro strip in the way that it adheres to RAD51, that is, through a large number of contacts that are relatively independent from one another. BRC repeats differing widely from the consensus may therefore retain a residual affinity towards RAD51, as the elimination of one or a few contact points would weaken binding, without abolishing it altogether. The BRC motif might have arisen as a molecular frame suitable for the evolution of amino acid sequences with a wide range of affinities to RAD51, with potential implications for the regulation of RAD51 function by BRCA2.

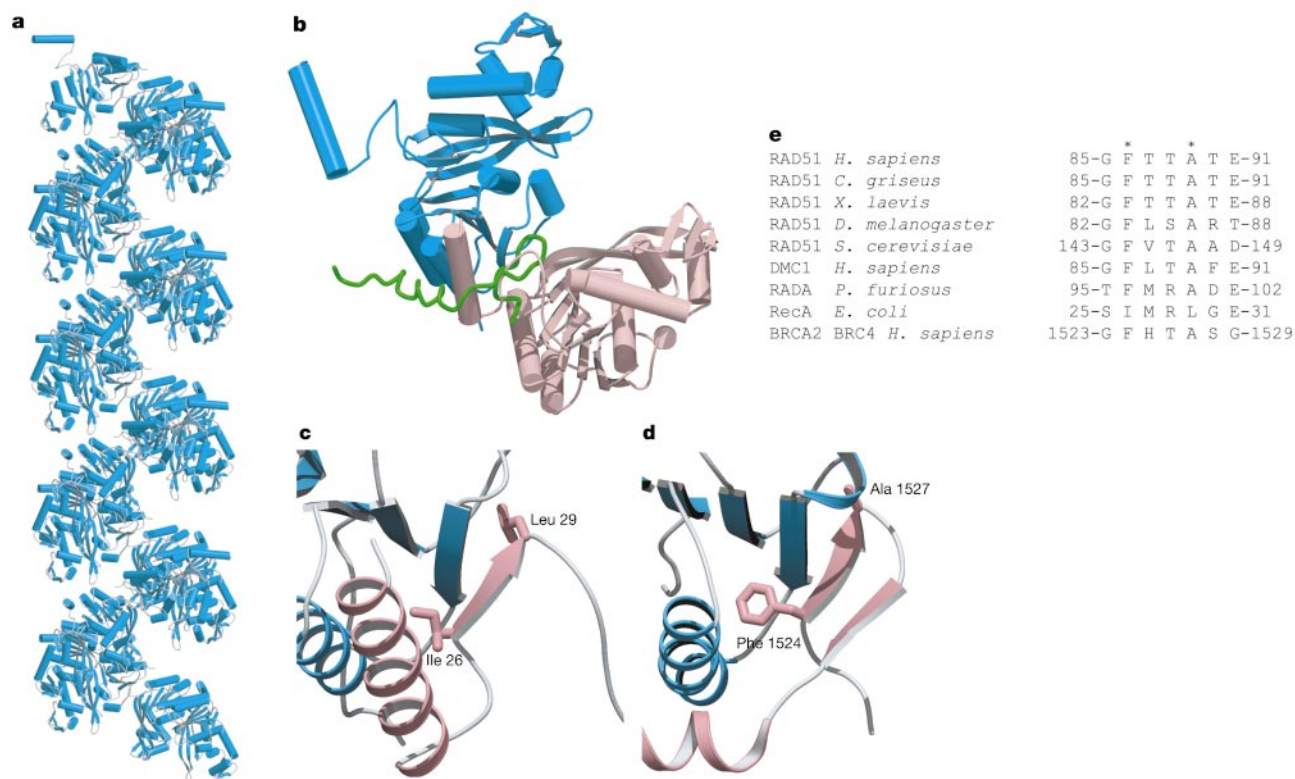


Figure 3 BRCA2 inhibits RAD51 filament formation. **a**, The crystallographic RecA filament. **b**, Superposition of the RAD51–BRC4 complex on a subunit of the crystallographic RecA filament (omitting RAD51 for clarity) positions the BRC repeat (green) at the interface between adjacent RecA subunits (in pink and blue). **c**, Detail of the subunit interface in the crystallographic RecA filament. The RecA sequence 26–IMRL–29 mediates polymerization by antiparallel β -strand pairing. Residues Ile 26 and Leu 29 represent points of hydrophobic contact between subunits. **d**, Detail of the interface between RAD51 (blue) and BRC4 (pink). The BRCA2 sequence 1524–FHTA–1527

interacts with RAD51 through antiparallel β -strand pairing. Residues Phe 1524 and Ala 1527 contact RAD51 hydrophobically. **e**, Evolutionary conservation of amino acids predicted to mediate filament formation in RecA-like recombinases. The relevant BRC4 sequence is also aligned. An asterisk marks the positions of two critical hydrophobic residues. *H. sapiens*, *Homo sapiens*; *C. griseus*, *Cricetulus griseus*; *X. laevis*, *Xenopus laevis*; *D. melanogaster*, *Drosophila melanogaster*; *S. cerevisiae*, *Saccharomyces cerevisiae*; *P. furiosus*, *Pyrococcus furiosus*; *E. coli*, *Escherichia coli*.

Regulation of RAD51 function by BRCA2

RAD51 forms helical nucleoprotein filaments on DNA substrates that catalyse pairing and strand exchange between homologous DNA molecules, an essential step in homologous recombination^{21–24}. Available biological data suggest that BRCA2 regulates RAD51 activity by controlling its aggregation state. *In vitro*, incubation with BRC repeat peptides maintains RAD51 in a monomeric form, thus removing its spontaneous tendency to oligomerize and form nucleoprotein filaments¹⁷. *In vivo*, RAD51 accumulates into nuclear foci in dividing cells, before or after exposure to DNA-damaging agents^{7,25,26}, and focus formation is suppressed by overexpression of BRC repeats⁷.

The structural basis for filament formation by RAD51 is not known. To gain an insight into the mechanism used by BRCA2 to regulate RAD51 filament formation, we analysed the RAD51–BRCA2 interaction in the context of the crystallographic RecA filament (Fig. 3a)¹⁸. In the crystal, the RecA molecules pack into a helical structure that closely resembles that of RecA/DNA filaments determined by electron microscopy¹⁸. Overlaying the RAD51–BRCA2 complex on RecA places the BRC4 β -hairpin at the interface between two adjacent RecA molecules of the crystallographic filament (Fig. 3b). Notably, BRC4 residues 1523–GFHTASG–1529 superimpose closely onto the RecA sequence 25–SIMRLGE–31, which forms part of the interface between RecA subunits. RecA residues 27–MRL–29 add an anti-parallel β -strand to the central β -sheet of a neighbouring RecA molecule, in an identical fashion to the interaction of BRC4 residues 1525–HTA–1527 with RAD51 in the RAD51–BRCA2 complex. Furthermore, RecA residues Ile 26 and Leu 29 make comparable hydrophobic contacts to those made by Phe 1524 and Ala 1527 of BRC4 with RAD51 (Fig. 3c, d).

The superposition analysis provides a compelling clue concerning the mechanism adopted by BRCA2 to regulate RAD51 function; that is, BRCA2 binding prevents formation of the nucleoprotein filament by interfering with a crucial contact between RAD51 subunits, and the specific role of the BRC repeats is to mimic the conformation of the RAD51 segment involved in such contact. One prediction of our proposed mechanism is that sequence similarity should be found between the BRC motif and the region of the RAD51 sequence with a putative role in oligomerization analogous to that of RecA sequence 25–SIMRLGE–31. Inspection of the RAD51

sequence for short motifs resembling the BRC consensus GFXTASG motif (where X is any amino acid) identifies the highly conserved sequence 85–GFTTATE–91 in the RAD51 linker between the N-terminal domain and the catalytic core (Fig. 3e).

To test the model of RAD51 oligomerization suggested by our crystallographic analysis, we constructed mutant RAD51 molecules in which amino acids Phe 86 and Ala 89 within the sequence 85–GFTTATE–91 were replaced by glutamic acid. Each of these changes should eliminate a critical hydrophobic contact at the RAD51 subunit interface and therefore abolish or significantly decrease the ability of RAD51 to form filaments. The RAD51 mutants were fused at their N termini to a green fluorescent protein (GFP) reporter before transfection into human cell lines. We find

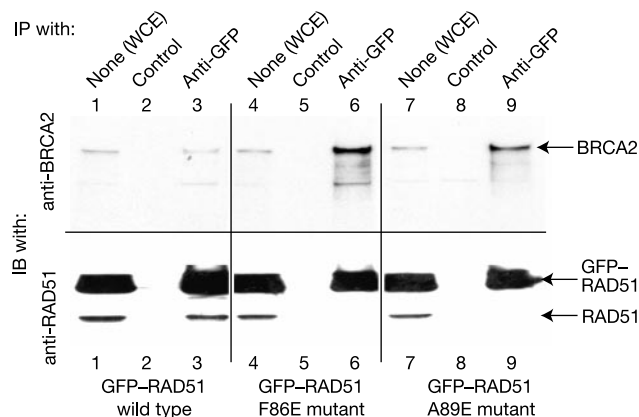


Figure 4 Mutational analysis of RAD51's self-association and BRCA2 binding. Analyses show cells transfected with wild-type (WT) GFP–RAD51 (lanes 1–3), GFP–RAD51 F86E mutant (lanes 4–6) or GFP–RAD51 A89E mutant (lanes 7–9). Lanes 1, 4 and 7 show BRCA2, GFP–RAD51 and endogenous RAD51 expression (arrows) in whole-cell extracts (WCE) from 10^5 cells. Immunoprecipitation (IP) of wild-type GFP–RAD51 from 3×10^6 cells with an anti-GFP antiserum shows its association with BRCA2 and endogenous RAD51 (lane 3). In contrast, the GFP–RAD51 F86E or A89E mutants fail to bind endogenous RAD51, although they can still associate with BRCA2 (lanes 6 and 9). Lanes 2, 5 and 8 show negative control immunoprecipitations.

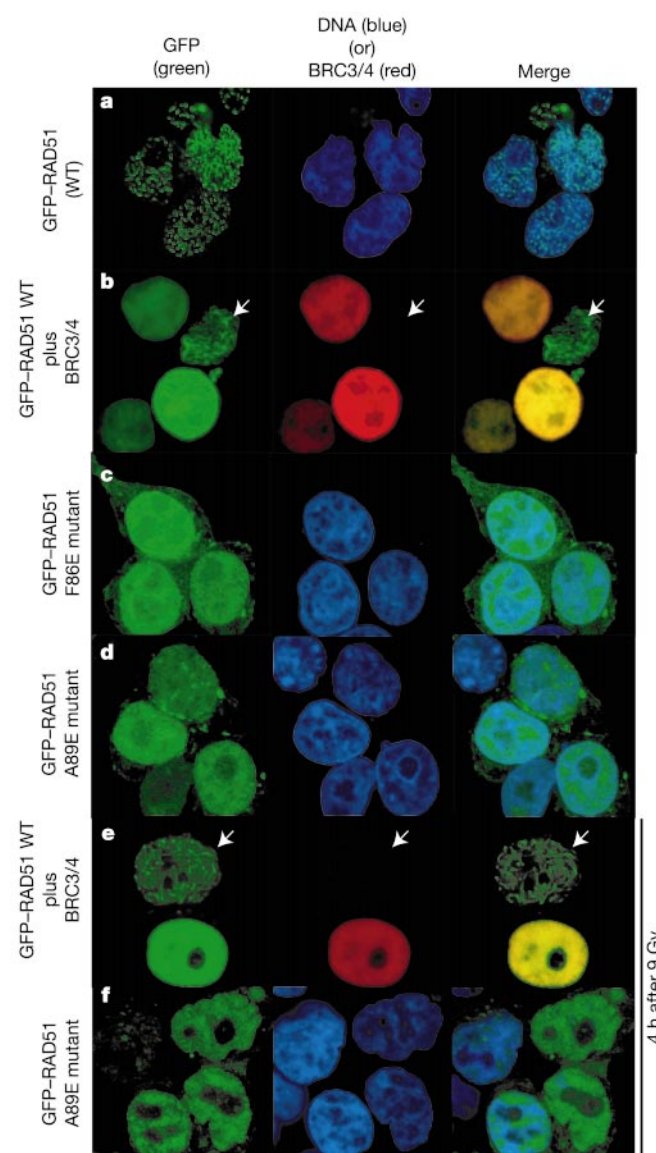


Figure 5 Mutational analysis of RAD51 focus formation. **a**, Wild-type GFP–RAD51 accumulates in nuclear foci. Nuclei are shown in blue pseudo-colour. **b**, Co-expression of BRC3 and BRC4 (red staining) prevents wild-type GFP–RAD51 focus formation. The green cell (white arrow) expresses GFP–RAD51 alone, and retains foci, providing an internal control. **c**, **d**, GFP-tagged RAD51 mutants do not accumulate in foci. **e**, Irradiation-induced wild-type GFP–RAD51 focus formation 4 h after exposure to 9 Gy is inhibited by co-expression of BRC3/4 (red). Focus formation occurs in an adjacent cell (white arrow) expressing GFP–RAD51 alone. **f**, GFP–RAD51 A89E mutants fail to accumulate in foci even after exposure to irradiation.

that a wild-type GFP–RAD51 fusion co-localizes with endogenously expressed RAD51, and can restore the viability of cells in which endogenous RAD51 has been eliminated by targeted gene disruption (D.S.Y., E. Sonoda, S. Takeda and A.V., manuscript in preparation).

The GFP–RAD51 wild type, GFP–RAD51 F86E and GFP–RAD51 A89E proteins are detected at equivalent levels in transfected cells, and their expression does not alter levels of endogenous RAD51 (Fig. 4, lanes 1, 4 and 7). Endogenous RAD51 was co-immunoprecipitated with wild-type GFP–RAD51 using anti-GFP (Fig. 4, lane 3), confirming the ability of GFP–RAD51 to form oligomers. In contrast, GFP–RAD51 F86E and GFP–RAD51 A89E failed to co-immunoprecipitate with endogenous RAD51, although their ability to bind BRCA2 remained intact (Fig. 4, lanes 6 and 9). The integrity of Phe 86 and Ala 89 is therefore essential for the ability of RAD51 to self-associate, in accordance with our prediction.

Next, we examined the impact of the F86E and A89E mutations on the ability of RAD51 to form nuclear foci. As previously observed for endogenous RAD51 (refs 25, 26), wild-type GFP–RAD51 accumulates in nuclear foci in dividing cells (Fig. 5a). Formation of the foci is not detected when a BRCA2 polypeptide spanning BRC repeats 3 and 4 is co-expressed in the same cells (Fig. 5b); a diffuse nuclear localization of wild-type GFP–RAD51 is observed instead, reminiscent of the distribution of GFP alone. The simplest explanation for the suppression of focus formation by the BRC repeats, in view of their known property to dissolve RAD51 aggregates *in vitro*, is that accumulation of RAD51 in the foci depends on its ability to oligomerize. We therefore predicted that the RAD51 mutations F86E and A89E should have a similar inhibitory effect. Indeed, we find that, when expressed in cells, GFP–RAD51 F86E (Fig. 5c) and GFP–RAD51 A89E (Fig. 5d) fail to form foci and are distributed diffusely throughout the nucleus. Exposure to ionizing radiation also fails to induce focus formation by GFP–RAD51 A89E (Fig. 5f).

Implications

Our crystallographic analysis of a human RAD51–BRCA2 complex has demonstrated that the conserved sequence motif 1523–GFHTASG–1529 of BRC4 structurally mimics the 25–SIMRLGE–31 sequence found at the interface between subunits in the bacterial RecA filament. Acting on this observation, we have produced evidence that RAD51 oligomerizes through a similar interface during nucleoprotein filament formation. First, the human RAD51–sequence 85–GFTTATE–91, located between the N-terminal and the ATPase domains, closely resembles the BRC consensus sequence GFXTASG. Second, replacement of Phe 86 or Ala 89 with glutamic acid creates RAD51 mutants that are no longer capable of self-association or nuclear focus formation when expressed in mammalian cells. Thus, BRCA2 interacts with RAD51 by mimicking a structural motif that enables RAD51 to oligomerize, preventing its incorporation into nucleoprotein filaments. Notably, the sequence that mediates the association between subunits in the RAD51 filament is conserved in the bacterial RecA, archaeal RADA and DMC1 protein families (Fig. 3e), thus pointing to a common evolutionary origin for the mechanism of nucleoprotein filament formation, a crucial intermediate in genetic recombination.

Structural analysis of cancer-associated mutations affecting the BRC repeats reveals that weakening RAD51 affinity in just one repeat is enough to increase breast cancer susceptibility. Why should the remaining seven BRC repeats not suffice to preserve function? One explanation is that the eight repeats cooperate as a RAD51-binding module whose overall topology is critical for function. By simultaneously interacting by means of successive BRC repeats, BRCA2 may therefore help to order the spatial distribution of RAD51 molecules when they are loaded onto DNA during nucleoprotein filament formation, or during removal from established filaments⁹. Alterations that diminish the RAD51 binding capacity of just one of the eight BRC repeats could perturb such functions by

interfering with spatial relationships between RAD51 molecules bound to BRCA2.

Regulation of RAD51 function by BRCA2 may also be modulated by physiological modifications^{9,17}. Phosphorylation of Thr 1526 would probably decrease RAD51-binding affinity by destabilizing the BRC repeat conformation, whereas phosphorylation of Ser 1528 or Ser 1538 would disrupt polar contacts with Asp 187 or Glu 213, respectively, in RAD51. The strong conservation of solvent-exposed lysine at position 1541 raises the possibility that it could be a target for covalent modifications. Thus, our analysis identifies several BRC repeat residues whose modification is predicted to affect complex formation, indicating how the pathways that signal DNA damage, blocked replication or cell cycle progression could regulate the RAD51–BRCA2 interaction.

Our findings provide a structural blueprint that may be useful in compound design. Our work shows that the RAD51–BRCA2 interaction will be particularly vulnerable to small-molecule inhibitors because it critically depends on spatially constrained hydrophobic contacts involving three BRCA2 residues, Phe 1524, Phe 1546 and Ala 1527, conserved among different BRC repeats. Because BRCA2 and RAD51 participate in the repair of DNA breakage during DNA replication^{8,9}, such inhibitors may prove useful adjuncts to radiation therapy or anti-cancer drugs that induce DNA damage or interfere with replication. □

Methods

Protein expression and purification

The BRCA2 BRC type 4 sequence (amino acids 1517 to 1551) was linked in the pGAT3 vector²⁷ to the N terminus of a RAD51 sequence spanning the RecA-homology domain (Ser 97 to the natural C terminus) by means of the flexible polypeptide sequence (Thr–Gly–Ser)₃Met–Gly, designed to allow for unrestrained interaction between the BRC motif and RAD51. Procedures for the expression and purification of the fusion protein are detailed in Supplementary Information.

Protein crystallization

Crystals of the BRC4–RAD51 fusion protein were grown in hanging drops by the vapour diffusion method. Drops were prepared by mixing equal volumes of protein and a 25% ethylene glycol reservoir solution. Crystals suitable for X-ray diffraction analysis grew at 18 °C within a few days to a maximum size of approximately 300 × 100 × 100 µm. The crystals belong to the space group $P2_12_12_1$ ($a = 57.30$ Å, $b = 59.14$ Å, $c = 77.20$ Å), with one BRC4–RAD51 fusion protein in the asymmetric unit.

Structure determination and refinement

The structure of the BRC4–RAD51 fusion protein complex was determined using phasing information from SIRAS and MAD experiments. An initial screening by native gel electrophoresis²⁸ identified KAu(CN)₂ as a potential heavy atom derivative. X-ray data from a native crystal soaked in 0.5 mM KAu(CN)₂ for 16 h were collected to 2 Å resolution. The position of the single gold site was readily determined using direct methods as implemented in the Shake 'N' Bake program²⁹. An initial set of phases was calculated with SHARP³⁰ and improved by the solvent modification routine available within the program. The resulting set of phases were further refined with ARP/WARP³¹, which successfully traced the entire chain of the BRC4 repeat and most of the RAD51 ATPase domain. We also prepared selenomethionine-substituted protein that crystallized under the same conditions as the native material. The selenomethionine-containing crystals were used to collect a two-wavelength MAD data set (peak and high-energy remote at the Se K-edge) at station ID29 of the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. The MAD phases were of excellent quality and allowed us to extend the resolution of the diffraction data to 1.7 Å and to considerably improve our model.

Crystallographic refinement was performed using the programs REFMAC³² and CNS³³. The refined model comprises 1,919 protein atoms, 239 water molecules and 4 ethylene glycol molecules. One magnesium ion and one chloride ion were also included in the final model to explain two strong, positive $F_o - F_c$ difference peaks, located at the C terminus of the short helix in the initial strand–helix–strand motif, and at the N terminus of helix A1. A total of 237 amino acid residues (98.8 %) are in the core region of the Ramachandran plot, 3 in the generously allowed region (1.2 %) and none in the disallowed region. RAD51 residues 97, 230–236 (loop L1 between β -strand B4 and helix H5), 268–292 (loop L2 between strands B5 and B6) and BRCA2 BRC4 residues 1517–1518 are not visible in the electron density map and are presumably disordered. The linker joining the BRC repeat to RAD51 is also not detectable in the map, with the exception of the initial Thr–Gly–Ser triplet. The quality of the map for the RAD51 region between strands B7 and B8 (residues 316–321) is poor, indicating that they are partially disordered in the crystals; the conformation of the polypeptide chain for this loop must therefore be considered tentative. Surface area accessibility calculations were carried out in CNS. Figures were prepared with Molscript³⁴ and Raster3D³⁵.

Analysis of RAD51 mutants

Phe 86 Glu or Ala 89 Glu mutations were introduced by site-directed mutagenesis using the QuickChange system (Stratagene) into a complementary DNA construct encoding the wild-type RAD51–GFP fusion in pEGFP-C1 (Clontech). The sequence encoding BRC3 and BRC4 from human BRCA2 (amino acids 1338–1617) was fused at its C terminus to three consensus nuclear localization signals in pEF-Myc-Nuc (Clontech). Constructs were verified by nucleotide sequencing. Observations were made 72–96 h after transfection into 293T cells, exposed where indicated to a measured dose of X-rays in an X-ray generator (Torres 150D, Astrophysics Research). For microscopy, we used a Zeiss LSM510 confocal system equipped with ZeissVision software. For biochemical experiments, whole-cell extracts were prepared from 10⁷ transfected cells in 1 ml of ice-cold lysis buffer (50 mM HEPES, pH 7.4, 100 mM NaCl, 0.5% Nonidet P-40, 10 mM EDTA, 20 mM β -glycerophosphate, 1 mM dithiothreitol, 1 mM phenylmethyl sulphonyl fluoride (PMSF) supplemented with protease inhibitors). Immunoprecipitations were from 0.3 ml of lysate using 4 μ g of control rabbit immunoglobulin- γ (IgG) (Sigma) or rabbit anti-GFP (AbCam). Immune complexes were captured on protein A-Sepharose (Amersham). Whole-cell extract and immunoprecipitates were resolved by 3–8% gradient SDS–polyacrylamide gel electrophoresis (PAGE), transferred to nylon membranes (Immobilon-P, Millipore) and blotted with antisera against either human BRCA2 (Ab1, Oncogene Research) or RAD51 (clone 3C10, Upstate Biotechnology).

Received 1 July; accepted 21 October 2002; doi:10.1038/nature01230.
Published online 10 November 2002.

1. Nathanson, K. N., Wooster, R. & Weber, B. L. Breast cancer genetics: what we know and what we need. *Nature Med.* **7**, 552–556 (2001).
2. Bork, P., Blomberg, N. & Nilges, M. Internal repeats in the BRCA2 protein sequence. *Nature Genet.* **13**, 22–23 (1996).
3. Bignell, G., Micklem, G., Stratton, M. R., Ashworth, A. & Wooster, R. The BRC repeats are conserved in mammalian BRCA2 proteins. *Hum. Mol. Genet.* **6**, 53–58 (1997).
4. Warren, M. *et al.* Structural analysis of the chicken BRCA2 gene facilitates identification of functional domains and disease causing mutations. *Hum. Mol. Genet.* **11**, 841–851 (2002).
5. Wong, A. K. C., Pero, R., Ormonde, P. A., Tavtigian, S. V. & Bartel, P. L. RAD51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene *brca2*. *J. Biol. Chem.* **272**, 31941–31944 (1997).
6. Chen, P. L. *et al.* The BRC repeats in BRCA2 are critical for RAD51 binding and resistance to methyl methanesulphonate treatment. *Proc. Natl Acad. Sci. USA* **95**, 5287–5292 (1998).
7. Chen, C. F., Chen, P. L., Zhong, Q., Sharp, Z. D. & Lee, W. H. Expression of BRC repeats in breast cancer cells disrupts the BRCA2–Rad51 complex and leads to radiation hypersensitivity and loss of G(2)/M checkpoint control. *J. Biol. Chem.* **274**, 32931–32935 (1999).
8. Scully, R. & Livingston, D. M. In search of the tumour-suppressor functions of BRCA1 and BRCA2. *Nature* **408**, 429–432 (2000).
9. Venkitaraman, A. R. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* **108**, 171–182 (2002).
10. Chen, G., Yuan, S. S., Liu, W. *et al.* Radiation-induced assembly of Rad51 and Rad52 recombination complex requires ATM and c-Abl. *J. Biol. Chem.* **274**, 12748–12752 (1999).
11. Yu, V. P. C. C. *et al.* Gross chromosomal rearrangements and genetic exchange between non-homologous chromosomes following BRCA2 inactivation. *Genes Dev.* **14**, 1400–1406 (2000).
12. Moynahan, M. E., Pierce, A. J. & Jasin, M. BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol. Cell* **7**, 263–272 (2001).
13. Lim, D. S. & Hasty, P. A mutation in mouse *rad51* results in an early embryonic lethal that is suppressed by a mutation in *p53*. *Mol. Cell Biol.* **16**, 7133–7143 (1996).
14. Tsuzuki, T. *et al.* Targeted disruption of the *Rad51* gene leads to lethality in embryonic mice. *Proc. Natl Acad. Sci. USA* **93**, 6236–6240 (1996).
15. Patel, K. J. *et al.* Involvement of Brca2 in DNA repair. *Mol. Cell* **1**, 347–357 (1998).
16. Tutt, A. *et al.* Absence of *brca2* causes genome instability by chromosome breakage and loss associated

with centrosome amplification. *Curr. Biol.* **9**, 1107–1110 (1999).

17. Davies, A. A. *et al.* Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol. Cell* **7**, 273–282 (2001).
18. Story, R. M., Weber, I. T. & Steitz, T. A. The structure of the *E. coli* recA protein monomer and polymer. *Nature* **355**, 318–325 (1992).
19. Sibanda, B. L. & Thornton, J. M. Beta-hairpin families in globular proteins. *Nature* **316**, 170–174 (1985).
20. Story, R. M. & Steitz, T. A. Structure of the recA protein-ADP complex. *Nature* **355**, 374–376 (1992).
21. Sung, P. & Robberson, D. L. DNA strand exchange mediated by a RAD51–ssDNA nucleoprotein filament with polarity opposite to that of RecA. *Cell* **82**, 453–461 (1995).
22. Baumann, P., Benson, F. E. & West, S. C. Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions *in vitro*. *Cell* **87**, 757–766 (1996).
23. Ogawa, T., Yu, X., Shinohara, A. & Egelman, E. H. Similarity of the yeast RAD51 filament to the bacterial RecA filament. *Science* **259**, 1896–1899 (1993).
24. Yu, X., Jacobs, S. A., West, S. C., Ogawa, T. & Egelman, E. H. Domain structure and dynamics in the helical filaments formed by RecA and Rad51 on DNA. *Proc. Natl Acad. Sci. USA* **98**, 8419–8424 (2001).
25. Haaf, T., Golub, E. I., Reddy, G., Radding, C. M. & Ward, D. C. Nuclear foci of mammalian Rad51 recombination protein in somatic cells after DNA damage and its localization in synaptonemal complexes. *Proc. Natl Acad. Sci. USA* **92**, 2298–2302 (1995).
26. Scully, R. *et al.* Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell* **88**, 265–275 (1997).
27. Peranen, J., Rikonen, M., Hyvonen, M. & Kaariainen, L. T7 vectors with modified T7lac promoter for the expression of proteins in *E. coli*. *Anal. Biochem.* **236**, 371–373 (1996).
28. Boggon, T. J. & Shapiro, L. Screening for phasing atoms in protein crystallography. *Structure Fold Des.* **8**, R143–R149 (2000).
29. Deacon, A. M., Weeks, C. M., Miller, R. & Ealick, S. E. The Shake-and-Bake structure determination of triclinic lysozyme. *Proc. Natl Acad. Sci. USA* **95**, 9284–9289 (1998).
30. La Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement in the MIR and MAD methods. *Methods Enzymol.* **276**, 472–494 (1997).
31. Perrakis, A., Morris, R. & Lamzin, V. S. Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458–463 (1999).
32. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **54**, 905–921 (1997).
33. Bruenger, A. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
34. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
35. Merritt, E. A. & Bacon, D. J. Raster3D photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524 (1997).

Supplementary Information accompanies the paper on *Nature's* website
(<http://www.nature.com/nature>).

Acknowledgements We thank A. Gupta for early work on the purification of the RAD51–BRCA2 complex; A. Thompson for technical assistance at the ESRF beamline ID29; M. Symmons for help with the figures; and R. Laskey for comments on this manuscript. This research was supported in the laboratory of T.L.B. by the Wellcome Trust, and in the laboratory of A.R.V. by the Medical Research Council and Cancer Research UK.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.R.V. (e-mail: arv22@cam.ac.uk). The coordinates have been deposited in the Protein Data Bank under accession code 1n0w.

The flux of small near-Earth objects colliding with the Earth

P. Brown^{*†}, R. E. Spalding[‡], D. O. ReVelle[†], E. Tagliaferri[§] & S. P. Worden^{||}

^{*} Department of Physics and Astronomy, University of Western Ontario, London, Ontario N6A 3K7, Canada

[†] Los Alamos National Laboratory, PO Box 1663, MS J577, Los Alamos, New Mexico 87545, USA

[‡] Sandia National Laboratory, Org. 5740, MS 0973, PO Box 5800, Albuquerque, New Mexico 87185, USA

[§] ET Space Systems, 5990 Worth Way, Camarillo, California 93012, USA

^{||} Directorate of Operations, United States Space Command, Peterson Air Force Base, Colorado Springs, Colorado 80914-3190, USA

Asteroids with diameters smaller than ~50–100 m that collide with the Earth usually do not hit the ground as a single body; rather, they detonate in the atmosphere¹. These small objects can still cause considerable damage, such as occurred near Tunguska², Siberia, in 1908. The flux of small bodies is poorly constrained, however, in part because ground-based observational searches pursue strategies that lead them preferentially to find larger objects³. A Tunguska-class event—the energy of which we take to be equivalent to 10 megatons of TNT—was previously estimated to occur every 200–300 years, with the largest annual airburst calculated to be ~20 kilotons (kton) TNT equivalent (ref. 4). Here we report satellite records of bolide detonations in the atmosphere over the past 8.5 years. We find that the flux of objects in the 1–10-m size range has the same power-law distribution as bodies with diameters >50 m. From this we estimate that the Earth is hit on average annually by an object with ~5 kton equivalent energy, and that Tunguska-like events occur about once every 1,000 years.

Our data are based on observations made by United States Department of Defense and Department of Energy space-based systems in geostationary orbits. These systems are designed to detect the signature of nuclear explosions and other objects of military interest on or above Earth's surface⁵. For the past eight years, data from optical events registered as probable bolides have been recorded and saved for later examination. In total, 300 events from February 1994 to September 2002 have been designated as probable bolides, and a time–intensity plot for each event recorded. From these time–intensity plots, the peak radiated power (maximum brightness) and the integrated energy were determined. Figure 1 shows an example of the original data for a large bolide recorded on 6 June 2002.

From this data set of recorded optical flashes, we obtain a number distribution of the optical energies of the 300 bolides in our sample. The total optical energy is expected to be an indicator of the original kinetic energy of the meteoroid⁶. We corrected the number distribution based on the percentage coverage of the Earth's surface, which varied from 60% to 80% throughout the period of study.

To convert these raw optical energy estimates to an equivalent source energy requires knowledge of several poorly known parameters. First, the spectral energy distribution at the source needs to be defined to determine the true radiated energy detected by the sensor. This process requires knowledge of the spectrum of the bolide. Previous workers have assumed a 6,000 K blackbody distribution⁵, but the dominant line emission from bright fireballs observed from the ground suggests that this assumed distribution might be a poor approximation⁷. In the absence of better data for these larger fireballs, we continue for consistency to use the 6,000 K blackbody assumption as others have also done⁸. As the peak in sensitivity for the optical sensors is near this temperature, our

corrections to the energy estimates will be more conservative than would be the case if the fireball temperatures were closer to 4,500 K (ref. 9). We expect the 6,000 K temperature assumption to produce less than ~30% uncertainty on the final flux numbers.

Second, a much greater unknown is establishing the fraction of total initial kinetic energy transformed into light production. This integral luminous efficiency (τ_i ; ref. 10) has been measured to be of the order 3% for photographically measured fireballs with energies between 10^{-5} and 10^{-1} kton (median of the sample, 7×10^{-4} kton)⁷. Theoretical models⁸ suggest that this value should be 10–15% for bolides of chondritic composition with energies between 10^{-1} and 10 kton, with the efficiency increasing as energy increases. More generally, we expect that the integral efficiency will be a complex function of velocity, mass, entry angle, body shape and composition, as well as porosity¹¹, differing from one bolide to the next. This effect is reflected in the wide variability observed in the analysis of high-precision ground-based bolide data where an order-of-magnitude, systematic variation in luminous efficiency between strong/chondritic bodies (highest values) to weak bodies (lowest) is found¹². No observational constraints for this value as a function of energy have been previously published for objects in our size range.

In an attempt to derive an empirical fit for τ_i for our population, we have examined those fireball events recorded by optical sensors and for which an independent estimate of the energy exists. A total of 13 events have optical data and independent energy estimates (Table 1).

The fit between optical energy and total calibrated energy (Fig. 2) leads to:

$$\tau_i = (0.1212 \pm 0.0043)E_o^{0.115 \pm 0.075} \quad (1)$$

where E_o is the optical energy (in kilotons TNT equivalent, 4.185×10^{12} J) measured by the space-based sensors assuming a 6,000 K blackbody emission from the bolides.

Applying this form of the integral luminous efficiency under our 6,000 K blackbody assumption to the bolide data set collected over 8.5 years, we find the flux as a function of impact energy (Fig. 3; see figure legend for details). We also convert these total impact energies into equivalent diameter of colliding body, assuming a mean impact velocity with the Earth of 20.3 km s^{-1} (W. Bottke, personal communication) and a bulk density of $3,000 \text{ kg m}^{-3}$, an appropriate compromise between ordinary chondritic densities (average near $3,400 \text{ kg m}^{-3}$) and primitive (CI/CM) carbonaceous chondrite densities (average near $2,600 \text{ kg m}^{-3}$)¹³. We note that below $\sim 10^{-1}$ kton, a roll-off in the cumulative number of bolides detected by the satellite is observed as the system limiting sensitivity is approached.

The power-law fit to the satellite data shown in Fig. 3 can also be

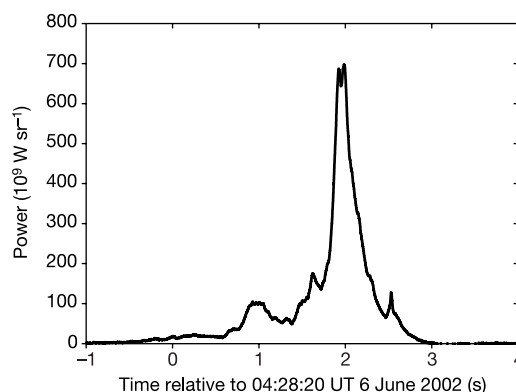


Figure 1 Optical light curve from the 6 June 2002 bolide over the Mediterranean Sea. The estimated energy for this event (on the basis of infrasonic/acoustic data) was 26 kton TNT equivalent.

Table 1 Details of calibrated bolides shown in Fig. 2

Bolide date	Time (UT)	Location (lat., long.)	Optical energy (kton)	Calibrated energy (kton)	Calibration method*
25/7/02	15:58	29° S, 47° E	0.060	0.60 ± 0.20	Infrasound
6/6/02	04:28	34° N, 21° E	0.908	25.77 ± 2.16	Infrasound
9/3/02	01:20	7° N, 147° W	0.053	1.05 ± 0.79	Infrasound
23/4/01	06:12	28° N, 134° W	1.100	0.63 ± 0.33	Infrasound
25/8/00	01:12	15° N, 106° W	0.337	2.35 ± 1.24	Infrasound
6/5/00	11:52	50° N, 18° E	0.006	0.09 ± 0.04	Infrasound/meteorites ²²
18/2/00	09:26	1° S, 109° E	0.860	3.89 ± 0.73	Infrasound
18/1/00	16:43	60° N, 135° W	0.263	1.66 ± 0.70	Infrasound/velocity/meteorites ²³
16/8/99	05:18	35° N, 107° W	0.009	0.18 ± 0.03	Infrasound
9/10/97	18:47	32° N, 106° W	0.045	0.29	Infrasound ²⁴
15/6/94	00:03	46° N, 73° W	0.003	0.03 ± 0.01	Meteorites/velocity ²⁵
7/5/91	23:04	50° N, 15° E	0.012	0.19	Photos/spectra/ velocity ²⁶
-	-	-	0.0007	0.0164	Photos/velocity ⁹

*Calibration methods include: 'infrasound', infrasound energy derived from period at maximum amplitude²⁰; 'meteorites', recovery mass estimate from cosmogenic nuclides; 'velocity', entry velocity and trajectory known; 'photos', multi-station photos of fireball; and 'spectra', spectral records of fireball. The final row of data represents the median values of the fireball events given in ref. 8.

represented by an equation of the form

$$\log N = a_0 - b_0 \log E \quad (2)$$

where N is the cumulative number of objects colliding with the Earth each year with energies of E (in kilotons) or greater; in this representation, $a_0 = 0.5677 \pm 0.015$, and $b_0 = 0.90 \pm 0.03$. We may recast this in terms of diameter of the colliding body using our assumptions above as:

$$\log N = c_0 - d_0 \log D \quad (3)$$

where D is the meteoroid diameter in metres, $c_0 = 1.568 \pm 0.03$, and $d_0 = 2.70 \pm 0.08$.

Comparing our result to debiased Lincoln Near Earth Asteroid Research (LINEAR) measurements¹⁴ as shown in Fig. 3, we see that

our power-law determination is in good agreement (within error) with the flux values estimated based on the small number (~ 30) of tens-of-metres class LINEAR discoveries. Influx estimates based on lunar crater¹⁵ counts extend down to meteoroids of several metres

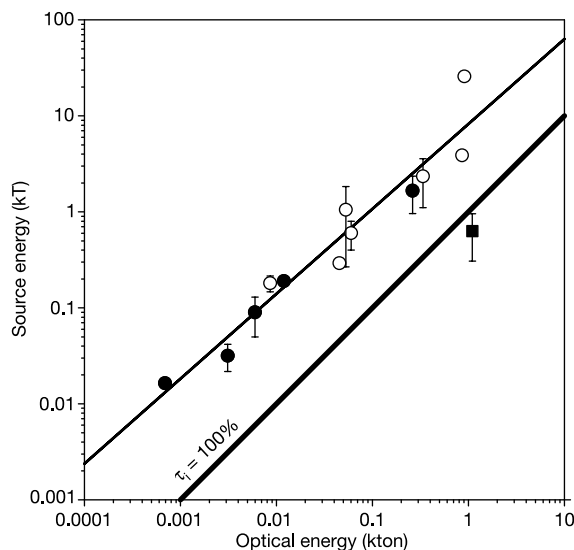


Figure 2 Bolide energy calibrations. These are based on simultaneous observations by optical sensors (or equivalent) and energy estimated from another technique. The optical energy estimates are made assuming a 6,000 K blackbody. Open circles, events that have also been detected infrasonic/acoustically and for which a source energy was computed using the calibrated relation between energy and period at maximum signal amplitude²⁰. Only events with periods greater than two seconds were used, to avoid issues of Doppler wind changes at small signal periods. Filled circles, events that have higher-precision source energy calibrations owing to recovery of meteorites and/or availability of information on atmospheric velocity—see Table 1 for details. The filled square represents the 23 April 2001 fireball event, which is particularly unusual²⁷ and was excluded from our analysis. The upper solid line shows the fit to the data represented by open and filled circles. The lower, thicker, solid line represents objects with integral luminous efficiencies of 100%; that is, those bodies to the right of this curve are unphysical within the context of our assumptions.

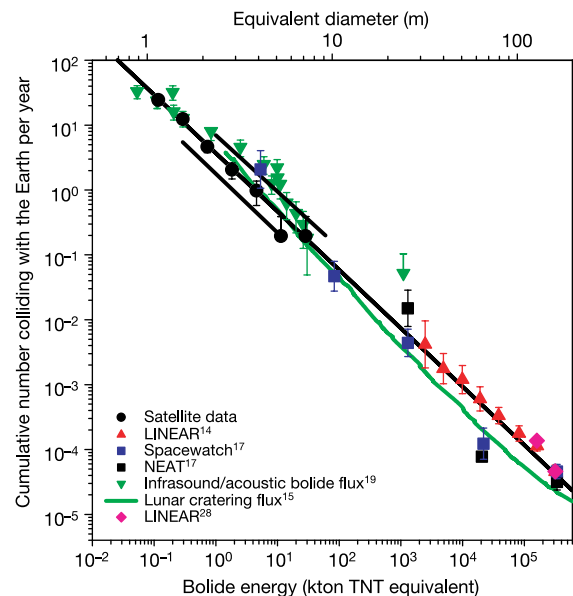


Figure 3 The flux of small near-Earth objects colliding with the Earth, for diameters less than 200 m. Filled circles, the satellite-determined flux based on 8.5 years of global observations (this work). The black best-fit line represents data for objects in excess of ~ 1 m in diameter where optical sensor data are most complete. Error bars represent counting statistics only. The thick black lines above and below the satellite flux represent the extreme limits, assuming integral luminous efficiencies for the entire population of 5% (top line) and 20% (bottom line). On the basis of the fit in Fig. 2 over our energy range of interest, τ_1 varies from 6% to 15%. From ground-truth calibrations and the analysis summarized in Fig. 2, we adopt 5% and 20% as the extreme probable limits for the average integral luminous efficiencies over the energy range represented by the lines. For comparison, we plot debiased estimates of the near-Earth asteroid impact frequency based on telescopic search data for smaller LINEAR discoveries (upward red triangles)¹⁴ and larger LINEAR discoveries²⁸ (pink diamonds), Near-Earth Asteroid Tracking (NEAT) (filled squares) and Spacewatch (blue squares)¹⁷ surveys, where diameters are determined assuming an albedo of 0.1. LINEAR values at the smaller sizes are normalized to the larger LINEAR discoveries²⁸. Energy for telescopic data is computed assuming a mean bulk density of $3,000 \text{ kg m}^{-3}$ and an average impact velocity of 20.3 km s^{-1} . The intrinsic impact frequency for these telescopic data was found by computing the average probability of impact for the 1,000 largest known near-Earth asteroids, and using this frequency ($2 \times 10^{-9} \text{ yr}^{-1}$) as the average for the entire population¹⁴. As well, the infrasonic/acoustically measured bolide flux (downward dark green triangles)¹⁹ and lunar crater counts assuming a geometric albedo of 0.25 (solid green line)¹⁵ are shown.

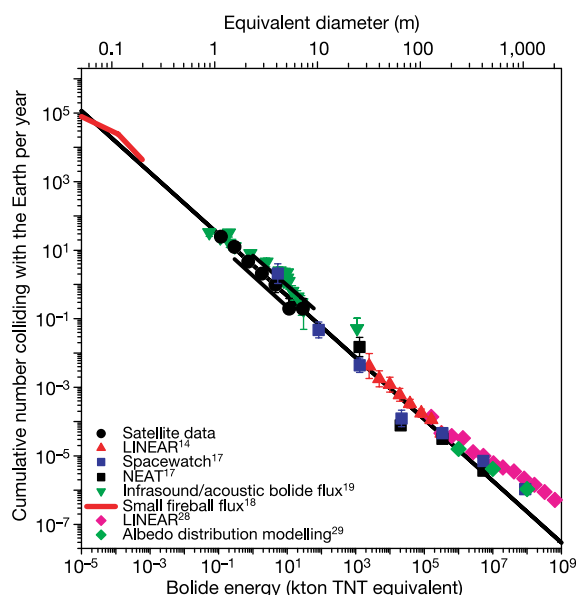


Figure 4 The flux of small near-Earth objects colliding with the Earth. Data are shown over a range of 14 magnitudes of energy. In addition to data shown in Fig. 3, this plot also shows the Earth collision hazard in the 1,000 Mton and larger energy range, based on modelling the albedo distribution of near-Earth asteroids²⁹. At smaller sizes, the power-law number distribution derived from a decade-long survey of ground-based observations of fireballs¹⁸ is indicated.

diameter, and we find a good fit with these data in the 3–10-m size range assuming an average geometric albedo of 0.25. We are unable to match the crater flux curve over the 3–10-m size range with our data for a typical main-belt mean albedo of 0.11 (near the average for C and S types in the main belt), further supporting the notion that small near-Earth asteroids have higher mean albedos than large main-belt asteroids¹⁶. We note, however, that the true geometric albedo distribution for the near-Earth asteroid population is uncertain.

The only telescopic flux measurement to overlap our population directly is that based on a debiasing of Spacewatch data¹⁷. In particular, the Spacewatch estimate (albeit with large error) of the flux of 4- and 10-m diameter bodies striking the Earth is in agreement with our satellite energy estimates within error limits.

We also find that extrapolation of our power law is in agreement with the flux of smaller bodies¹⁸ for energies between 10^{-5} kton and 10^{-3} kton (Fig. 4), these data being based on observations with ground-based cameras of fireballs having energies of $<10^{-3}$ kton. However, the influx measurements derived from infrasonic/acoustic measurements are slightly higher on average than the satellite estimates (notably between 1 kton and 10 kton), though in agreement near the upper end of our energy range where the infrasonic/acoustic set has the best overall global coverage. One possible reason for the difference is the small number of statistics associated with the infrasonic/acoustic data (only 19 events in total). Alternatively, a large range in meteoroid porosity may influence the optical light curve interpretation, potentially making an order-of-magnitude difference between the estimated and true energy. That the infrasonic/acoustic and satellite influx agree in the size range in which both are accurate (~ 7 –10 m) is further evidence that the most of the flux of bodies striking the Earth is asteroidal, rather than cometary, as only deeply penetrating bodies can generate infrasonic/acoustic shocks and thus be detected with these infrasonic/acoustic systems^{19,20}. Agreement in fit between our power-law extrapolation (solid line in Fig. 3) and LINEAR data, as well as with the cratering record and infrasonic/acoustic data at larger sizes, further suggests that objects of cometary origin make a small contribution to the flux

of 1–10-m bodies striking the Earth²¹.

Using the best fit of these satellite data and extrapolating the power law to higher energies, we find that the Earth is struck by an object with the energy of Tunguska (assumed to be 10 Mton) every $1,000^{+800}_{-200}$ years (with an allowed range from 400 to 1,800 years on the basis of our most extreme assumptions for luminous efficiency). We estimate that the Earth is on average struck annually by an object of energy ~ 5 kton (with a possible range of 2–10 kton), and struck each month by an object with 0.3 kton of energy. Every ten years, an object of energy ~ 50 kton strikes Earth. □

Received 9 October; accepted 21 October 2002; doi:10.1038/nature01238.

- Hills, J. G. & Goda, P. Damage from the impacts of small asteroids. *Planet. Space Sci.* **46**, 219–229 (1998).
- Sekanina, Z. Evidence for the asteroidal origin of the Tunguska object. *Planet. Space Sci.* **46**, 191–204 (1998).
- Morrison, D., Chapman, C. R. & Slovic, P. *Hazards due to Comets and Asteroids* (ed. Gehrels, T.) 59–91 (Univ. Arizona Press, Tucson, 1994).
- Shoemaker, E. M. Asteroid and comet bombardment of the Earth. *Annu. Rev. Earth Planet. Sci.* **11**, 461–494 (1983).
- Tagliaferri, E., Spalding, R., Jacobs, C., Worden, S. P. & Erlich, A. *Hazards due to Comets and Asteroids* (ed. Gehrels, T.) 199–221 (Univ. Arizona Press, Tucson, 1994).
- McCord, T. B. et al. Detection of a meteoroid entry into the Earth's atmosphere on February 1, 1994. *J. Geophys. Res.* **100**, 3245–3249 (1995).
- Cepelcha, Z. et al. Meteor phenomena and bodies. *Space Sci. Res.* **84**, 327–471 (1998).
- Nemtchinov, I. et al. Assessment of kinetic energy of meteoroids detected by satellite-based light sensors. *Icarus* **130**, 259–274 (1997).
- Cepelcha, Z., Spalding, R. E., Jacobs, C. & Tagliaferri, E. *Proc. SPIE* **2813**, 46–56 (1996).
- ReVelle, D. O. A predictive macroscopic integral radiation efficiency model. *J. Geophys. Res.* **85**, 1803–1808 (1980).
- ReVelle, D. O. *Meteoroids 2001—Conference Proceedings SP-495* (ed. Warmbein, B.) 513–519 (European Space Agency, Noordwijk, The Netherlands, 2001).
- ReVelle, D. O. & Cepelcha, Z. *Meteoroids 2001—Conference Proceedings SP-495* (ed. Warmbein, B.) 507–512 (European Space Agency, Noordwijk, The Netherlands, 2001).
- Britt, D. & Consolmagno, G. The porosity of dark meteorites and the structure of low-albedo asteroids. *Icarus* **146**, 213–219 (2000).
- Harris, A. W. in *Proc. Asteroids, Comets, Meteors 2002* (Berlin, in the press).
- Werner, S. C., Harris, A. W., Neukum, G. & Ivanov, B. A. NOTE: The near-Earth asteroid size-frequency distribution: a snapshot of the lunar impactor size-frequency distribution. *Icarus* **156**, 287–290 (2002).
- Harris, A. W. & Davies, J. K. Physical characteristics of near-Earth asteroids from thermal infrared spectrophotometry. *Icarus* **142**, 464–475 (1999).
- Rabinowitz, D., Helin, E., Lawrence, K. & Pravdo, S. A reduced estimate of the number of kilometre-sized near-Earth asteroids. *Nature* **403**, 165–166 (2000).
- Halliday, I., Griffin, A. A. & Blackwell, A. T. Detailed data for 259 fireballs from the Canadian camera network and inferences concerning the influx of large meteoroids. *Meteoritics* **31**, 185–217 (1996).
- ReVelle, D. O. *Meteoroids 2001—Conference Proceedings SP-495* (ed. Warmbein, B.) 483–498 (European Space Agency, Noordwijk, The Netherlands, 2001).
- ReVelle, D. O. Historical detection of atmospheric impacts by large bolides using acoustic-gravity waves. *Ann. NY Acad. Sci.* **822**, 284–302 (1997).
- Levison, H. F. et al. The mass disruption of Oort cloud comets. *Science* **296**, 2212–2215 (2002).
- Borovička, J. et al. The Moravka meteorite fall – III: Meteoroid initial size, history and composition. *Meteorit. Planet. Sci.* (submitted).
- Brown, P. et al. An entry model for the Tagish Lake fireball using seismic, satellite and infrasound records. *Meteorit. Planet. Sci.* **37**, 661–675 (2002).
- ReVelle, D. O., Whitaker, R. W., Armstrong, R. T. *Proc. SPIE* **3434**, 66–77 (1998).
- Brown, P. et al. The fall of the St Robert meteorite. *Meteorit. Planet. Sci.* **31**, 502–517 (1995).
- Borovička, J. & Spurný, P. Radiation study of two very bright terrestrial bolides and an application to the comet S-L 9 collision with Jupiter. *Icarus* **121**, 484–510 (1996).
- Brown, P., Whitaker, R. W., ReVelle, D. O. & Tagliaferri, E. Multi-station infrasonic observations of two large bolides: signal interpretation and implications for monitoring of atmospheric explosions. *Geophys. Res. Lett.* **29**, 1–4 (2002).
- Stewart, J. S. A near-Earth asteroid population estimate from the LINEAR survey. *Science* **294**, 1691–1693 (2001).
- Morbidelli, A., Jedicke, R., Bottke, W. F., Michel, P. & Tedesco, E. F. From magnitudes to diameters: The albedo distribution of near-Earth objects and the Earth collision hazard. *Icarus* **158**, 329–343 (2002).

Acknowledgements We thank the US Department of Defense for making these data available. We also thank A.W. Harris, J. Borovička and Z. Cepelcha for discussions; A.W. Harris for making available to us LINEAR debiased data before publication; and W. Bottke and R. Jedicke for comments and suggestions that improved an earlier version of this Letter. This work was supported in part by the Canada Research Chair program, and the Natural Sciences and Engineering Research Council of Canada. There is no implied endorsement by the US Department of Defense of factual accuracy or opinion in this paper.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.B. (e-mail: pbrown@uwo.ca).

Plutonium-based superconductivity with a transition temperature above 18 K

J. L. Sarrao*, L. A. Morales*, J. D. Thompson*, B. L. Scott*, G. R. Stewart*†, F. Wastin‡, J. Rebizant‡, P. Boulet‡, E. Colineau‡ & G. H. Lander*‡

* Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

† Department of Physics, University of Florida, Gainesville, Florida 32611, USA

‡ European Commission, JRC, Institute for Transuranium Elements, Postfach 2340, 76125 Karlsruhe, Germany

Plutonium is a metal of both technological relevance and fundamental scientific interest. Nevertheless, the electronic structure of plutonium, which directly influences its metallurgical properties¹, is poorly understood. For example, plutonium's 5*f* electrons are poised on the border between localized and itinerant, and their theoretical treatment pushes the limits of current electronic structure calculations². Here we extend the range of complexity exhibited by plutonium with the discovery of superconductivity in PuCoGa₅. We argue that the observed superconductivity results directly from plutonium's anomalous electronic properties and as such serves as a bridge between two classes of spin-fluctuation-mediated superconductors: the known heavy-fermion superconductors and the high-*T_c* copper oxides. We suggest that the mechanism of superconductivity is unconventional; seen in that context, the fact that the transition temperature, *T_c* ≈ 18.5 K, is an order of magnitude greater than the maximum seen in the U- and Ce-based heavy-fermion systems may be natural. The large critical current displayed by PuCoGa₅, which comes from radiation-induced self damage that creates pinning centres, would be of technological importance for applied superconductivity if the hazardous material plutonium were not a constituent.

Large, well-faceted single crystals of PuCoGa₅ were grown by combining Pu and Co with excess Ga in an alumina crucible. The crucible was encapsulated in an evacuated quartz ampoule, which was then heated to 1,100 °C and cooled over forty hours to 600 °C. At this point the excess flux was removed with a centrifuge. Single-crystal X-ray diffraction measurements reveal that the resulting material is tetragonal with room-temperature lattice constants *a* = 4.232 Å and *c* = 6.786 Å. This structure consists of alternating PuGa₃ and 'CoGa₂' layers and is identical to the HoCoGa₅ structure reported for uranium-based gallides with quite similar lattice constants³. This is also the structure type in which a family of unconventional Ce-based superconductors CeMIn₅ (where M is Co, Ir, Rh) crystallizes⁴. Magnetic susceptibility, specific heat, and electrical resistivity measurements on PuCoGa₅ were made as functions of temperature and magnetic field. Some of the results obtained at Los Alamos have been reproduced at Karlsruhe (using different starting ²³⁹Pu) by extracting small single-crystal platelets from arc-melted materials. The excellent reproducibility of the results from the two laboratories demonstrates the robust nature of the observed ground-state properties.

Low-temperature magnetic susceptibility and specific heat measurements on PuCoGa₅ are shown in Fig. 1. Zero-field cooled magnetization measured in a field of 10 Oe reveals a sharp diamagnetic transition at 18.5 K. At low temperature this signal corresponds to almost 100% of perfect diamagnetism. Heat capacity measurements confirm a bulk phase transition. The inferred Sommerfeld coefficient $\gamma = 77 \text{ mJ mol}^{-1} \text{ K}^{-2}$ for PuCoGa₅ is comparable to the value observed for δ -Pu ($\gamma = 50 \text{ mJ mol}^{-1} \text{ K}^{-2}$) (refs 5, 6) and is indicative of modest quasiparticle mass enhancement.

Taken together, these data provide unambiguous evidence for bulk superconductivity in PuCoGa₅, the first such observation in a plutonium compound. The approximately 0.25-K difference in *T_c* between the magnetic and heat capacity measurements provides a simple indication that the compound in question contains Pu: the *T_c* of PuCoGa₅ decreases at a rate of about 0.2 K per month, presumably as a result of radiation-induced self damage. As a result, non-simultaneous measurements yield slightly different *T_c* values. This effect is also evident in magnetization measurements. Measurements made on the same sample, but separated in time by approximately two months, reveal a decrease in *T_c* of 0.4 K.

Such a high *T_c* is unusual for an intermetallic compound (only a small number of intermetallics, led by MgB₂ with a *T_c* of 39 K (ref. 7), have *T_c* values exceeding 18 K). The upper critical field in PuCoGa₅ is correspondingly large. Figure 2 shows field-dependent resistivity data and the resulting upper critical field *H_{c2}*(*T*) phase diagram inferred from these data. In particular, we find an initial slope dH_{c2}/dT of -59 kOe K^{-1} . In the WHH approximation⁸, the orbital upper critical field $H_{c2}(0) = -0.69 T_c dH_{c2}/dT = 740 \text{ kOe}$. This estimate is quite large and exceeds the Pauli limit ($H_p = 18.6 (\text{kOe K}^{-1}) T_c = 340 \text{ kOe}$) (ref. 9). From this value of the orbital *H_{c2}*, we infer the Ginzburg–Landau coherence length $\xi_{GL} = [\Phi_0/2\pi H_{c2}(0)]^{0.5} = 2.1 \text{ nm}$. Further, assuming the Bardeen–Cooper–Schrieffer (BCS) coherence length $\xi_{BCS} \approx \xi_{GL}$ and that Pu is trivalent (see below), we estimate $\gamma \approx 58 \text{ mJ mol}^{-1} \text{ K}^{-2}$ in the free-electron approximation, consistent with the data of Fig. 1.

The lower inset of Fig. 2 presents a complete magnetization loop *M*(*H*) for PuCoGa₅ at 5 K. We estimate the lower critical field

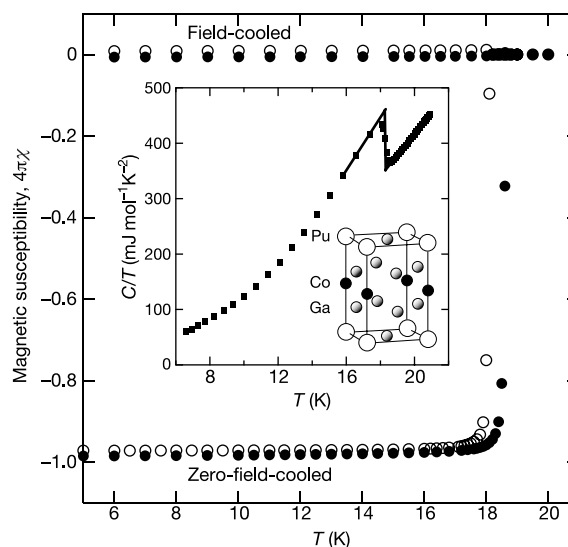


Figure 1 Crystal structure and evidence for superconductivity in PuCoGa₅. PuCoGa₅ crystallizes in the *P4/mmm* space group as follows: Pu, 1a, (0,0,0); Co, 1b, (0,0,0.5); Ga1, 1c, (0.5,0.5,0); Ga2, 4i, (0,0.5,0.312). Zero-field-cooled and field-cooled magnetic susceptibility of PuCoGa₅ were measured in 10 Oe for a fresh (filled symbols) and an aged (open symbols) single crystal. Magnetization measurements were performed in a Quantum Design superconducting quantum interference device (SQUID) magnetometer with the sample sealed in an alumina holder designed to minimize background signal and prevent spread of radioactive contamination. The inset shows heat capacity, plotted as heat capacity, *C*, divided by temperature versus temperature for PuCoGa₅. Heat capacity measurements were made in a Quantum Design Physical Property Measurement System on a 27-mg single crystal. Self-heating limited the lowest measurement temperature to 6.6 K, although the calorimeter reached a base temperature of 1.5 K. The jump in *C/T* at *T_c* corresponds to $\Delta C/T_c = 110 \pm 4 \text{ mJ mol}^{-1} \text{ K}^{-2}$. Assuming the BCS value for $\Delta C/\gamma T_c$ yields $\gamma = 77 \text{ mJ mol}^{-1} \text{ K}^{-2}$. Fitting the heat capacity data above *T_c* yields an estimate of the Debye temperature of 240 K.

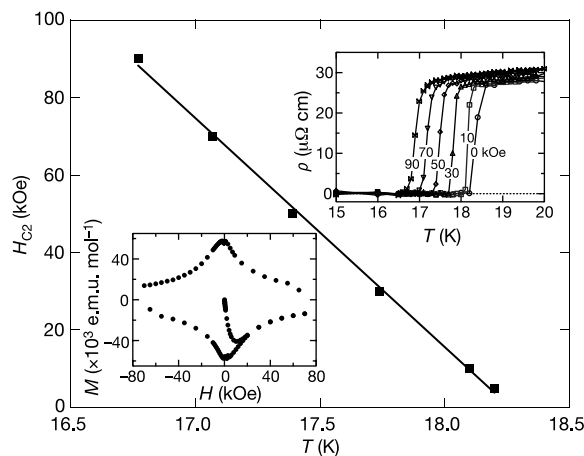


Figure 2 Upper critical field of PuCoGa₅ as a function of temperature. The upper inset shows the field-dependent resistivity data from which the field-temperature phase diagram was deduced. The lower inset shows a representative magnetization loop for PuCoGa₅, measured at 5 K.

$H_{c1} \approx 350$ Oe from the field at which M deviates from its initial linear H dependence. The approximate linearity of $M(H)$ to much higher fields is indicative of strong flux pinning and consistent with the absence of a strong Meissner effect in the field-cooled data of Fig. 1. The geometric mean of H_{c1} and H_{c2} gives the thermodynamic critical field $H_c \approx 16$ kOe. These values of H_{c2} , H_{c1} , and H_c indicate that PuCoGa₅ is a strongly type II superconductor with Ginzburg-Landau parameter $\kappa \approx 32$, and London penetration depth about 124 nm. From the condensation energy $-1/2N(0)\Delta^2(0) = H_c^2/8\pi\mu_0$ and the BCS result $\Delta(0) = 1.76kT_c$, we estimate $\gamma = 73$ mJ mol⁻¹ K⁻², again consistent with the data of Fig. 1, where $N(0)$ is the density of electronic states, Δ is the superconducting gap, μ_0 is permeability and k is Boltzmann's constant. Magnetization data also have been obtained for a larger, more well-shaped single crystal; however, the data are limited to $T > 0.9T_c$ owing to saturation of the SQUID detection system. A simple estimate of the critical current can be made from these higher temperature data using the Bean critical state model¹⁰. We find low-field values of $J_c > 10^4$ A cm⁻² for $T > 0.9T_c$. Such values are competitive with the best available applied superconductors. Further, in contrast to the degradation of T_c , J_c increases with time by a factor of nearly 2 over the same time period that T_c decreases by 0.4 K, evidently owing to the increasing number of radiation-induced pinning centres.

The superconducting properties of PuCoGa₅ are surprisingly well suited to the self-damage mechanism of Pu. The radioactive decay of ²³⁹Pu results in the formation of a high-energy alpha particle and a U nucleus. The principal damage is done by the U nucleus, which is displaced by approximately 12 nm and creates about 2,300 Frenkel pairs of vacancies and displaced interstitials distributed over a range of 7.5 nm (ref. 11). This damage is randomly distributed throughout the bulk of the material because it arises from spontaneous decay of Pu. Defects with spatial dimensions of the order of the superconducting coherence length create effective flux pinning and high critical currents (ref. 12).

A detailed understanding of what gives rise to such high-temperature superconductivity in PuCoGa₅ must await comprehensive studies of both its superconducting and normal states; however, available data are suggestive. Figure 3 shows magnetic susceptibility and electrical resistivity data over a broad temperature range for PuCoGa₅. The temperature dependence of the electrical resistivity is reminiscent of that of UMGa₅ (refs 3, 13) and suggests the presence

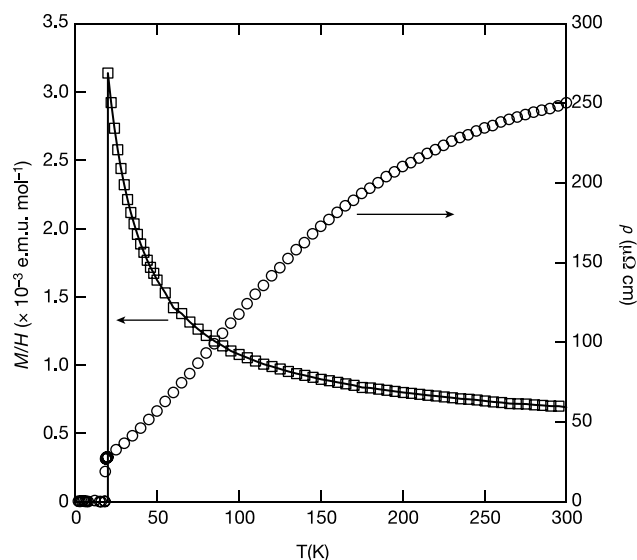


Figure 3 Normal-state properties of PuCoGa₅. The electrical resistivity ρ (circles) increases approximately as $T^{1.35}$ from just above T_c to 50 K. The magnetic susceptibility $\chi \equiv M/H$ (squares) as a function of temperature follows $\chi = \chi_0 + C/(T - \theta)$ with an effective moment $\mu = (8C)^{1/2} = 0.68\mu_B$ and an interaction temperature $\theta \approx -2$ K.

of spin-disorder scattering at high temperature. The temperature dependence of the magnetic susceptibility of PuCoGa₅ is indicative of local-moment behaviour close to that expected for Pu³⁺. A local-moment susceptibility is also found in CeCoIn₅ (ref. 4) but isostructural UCoGa₅ (ref. 3) displays temperature-independent paramagnetism consistent with itinerant f -electron behaviour. Taken together, these data suggest that the degree of $5f$ -electron localization in PuCoGa₅ falls between that of its Ce-based and U-based analogues. This conclusion is consistent with our several estimates of the low-temperature enhancement of electronic heat capacity and leads to the reasonable speculation that the superconductivity in PuCoGa₅ may be unconventional. Although this may be questioned, the alternative, 18-K phonon-mediated superconductivity in the presence of local-moment susceptibility, is equally challenging. Moreover, the isostructural UCoGa₅, which shows no sign of a local moment, is not a superconductor.

In a scenario of unconventional superconductivity, the nearly order-of-magnitude-higher T_c in PuCoGa₅ relative to that of CeCoIn₅ would be attributable to increased hybridization consistent with predictions for models of magnetically mediated superconductivity¹⁴. Further, the $5f$ electrons of the actinides are intermediate between the more localized $4f$ electrons of the rare earths and the itinerant d electrons of the transition metals. The combination of layered crystal structure with greater delocalization may suggest that the $T_c \approx 200$ mK observed in CeIn₃ (ref. 14), the three-dimensional, low- T_c analogue of CeCoIn₅ (ref. 4), is not unrelated to the $T_c \approx 20$ K of PuCoGa₅. As a result, the transuranics may represent a promising field for superconductivity, intermediate between the known heavy-fermion superconductors and the high- T_c copper oxides. □

Received 2 September; accepted 14 October 2002; doi:10.1038/nature01212.

- Hecker, S. S. The complex world of plutonium science. *MRS Bull.* **26**, 672–678 (2001).
- Savrasov, S. Y., Kotliar, G. & Abrahams, E. Correlated electrons in δ -plutonium within a dynamical mean-field picture. *Nature* **410**, 793–795 (2001).
- Grin, Yu. N., Rogl, P. & Hiebl, K. Structural chemistry and magnetic behavior of ternary uranium gallides U(Fe,Co,Ni,Ru,Rh,Pd,Os,Ir,Pt)-Ga₅. *J. Less Common Met.* **121**, 497–505 (1986).
- Petrovic, C. *et al.* Heavy-fermion superconductivity in CeCoIn₅ at 2.3 K. *J. Phys. Condens. Matt.* **13**, L337–L342 (2001).

5. Wick, O. J. (ed.) *Plutonium Handbook: a Guide to the Technology* 33–57 (American Nuclear Society, LaGrange Park, IL, 1980).
6. Stewart, G. R. & Elliott, R. O. *Actinides in Perspective Abstracts* Lawrence Berkeley Laboratory Report no. 12441 206–207 (Lawrence Berkeley National Laboratory, Berkeley, CA, 1981).
7. Nagamatsu, J., Nakagawa, N., Muranaka, T., Zenitani, Y. & Akimitsu, J. Superconductivity at 39 K in magnesium diboride. *Nature* **410**, 63–64 (2001).
8. Werthamer, N. R., Helfand, E. & Hohenberg, P. C. Temperature and purity dependence of the superconducting critical field, H_{c2} . III. Electron spin and spin-orbit effects. *Phys. Rev.* **147**, 295–302 (1966).
9. Clogston, A. M. Upper limit for the critical field in hard superconductors. *Phys. Rev. Lett.* **9**, 266 (1962).
10. Bean, C. P. Magnetization of hard superconductors. *Phys. Rev. Lett.* **8**, 250–253 (1962).
11. Wolfer, W. G. Radiation effects in plutonium. *Los Alamos Sci.* **26**, 274–285 (2000).
12. Campbell, A. M. & Evetts, J. E. Flux vortices and transport currents in type II superconductors. *Adv. Phys.* **21**, 199–427 (1972).
13. Tokiwa, Y. et al. Magnetic and Fermi surface properties of UPtGa₅. *J. Phys. Soc. Jpn* **71**, 845–851 (2002).
14. Mathur, N. D. et al. Magnetically mediated superconductivity in heavy fermion compounds. *Nature* **394**, 39–43 (1998).

Acknowledgements We thank Z. Fisk and C. Varma for discussions. Work at Los Alamos and Florida was performed under the auspices of the US Department of Energy.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.L.S. (e-mail: sarrao@lanl.gov).

Electric-field-induced capillary attraction between like-charged particles at liquid interfaces

M. G. Nikolaides^{*†}, A. R. Bausch^{*†}, M. F. Hsu^{*}, A. D. Dinsmore^{*†},
M. P. Brenner[‡], C. Gay[§] & D. A. Weitz^{*‡}

^{*} Department of Physics; and [‡] Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA
[§] CRPP-CNRS, Avenue Schweitzer, 33600 Pessac, France

Nanometre- and micrometre-sized charged particles at aqueous interfaces are typically stabilized by a repulsive Coulomb interaction. If one of the phases forming the interface is a nonpolar substance (such as air or oil) that cannot sustain a charge, the particles will exhibit long-ranged dipolar repulsion¹; if the interface area is confined, mutual repulsion between the particles can induce ordering² and even crystallization^{3,4}. However, particle ordering has also been observed in the absence of area confinement⁵, suggesting that like-charged particles at interfaces can also experience attractive interactions⁶. Interface deformations are known to cause capillary forces that attract neighbouring particles to each other, but a satisfying explanation for the origin of such distortions remains outstanding^{7,8}. Here we present quantitative measurements of attractive interactions between colloidal particles at an oil–water interface and show that the attraction can be explained by capillary forces that arise from a distortion of the interface shape that is due to electrostatic stresses caused by the particles' dipolar field. This explanation, which is consistent with all reports on interfacial particle ordering so far, also suggests that the attractive interactions might be controllable: by tuning the polarity of one of the interfacial fluids, it should be possible to adjust the electro-

static stresses of the system and hence the interparticle attractions.

The mismatch in dielectric constants of adjacent fluids can result in asymmetric charging of particles adsorbed at their interface; if one of the fluids is water, then the aqueous surface acquires a charge, which combines with the screening ions in the water to produce an effective dipole moment of the particle, leading to repulsion^{1,2,9}. By contrast, the origin of an attractive interaction is less well understood, even though interface deformation is known to give rise to capillary forces and a logarithmic attraction between neighbouring particles^{10–12}. For sufficiently large particles, the deformation is caused by gravity; the particle weight is balanced by surface tension, deforming the interface⁴. This results in the clumping of breakfast cereals at the surface of a bowl of milk, and has been harnessed to self-assemble millimetre-sized particles at interfaces¹³. However, the buoyancy mismatch of micrometre-sized colloidal particles is too small to deform the interface significantly and the resulting attractive energies are far smaller than thermal energy¹⁴.

An alternative candidate for the origin of the attraction is wetting of the particles, which also deforms the interface; however, a spherical particle positions itself exactly to achieve the equilibrium contact angle without distorting the interface unless the particles are constrained to a thin layer of fluid^{7,14}. Another candidate for the origin of the attraction is surface roughness, which might also deform the interface; however, the roughness required is far greater than that which typically exists on colloidal particles⁸. A final candidate for the origin of the attraction is thermal fluctuations; entropic interactions lower the free energy of the interface when two particles approach, resulting in a Casimir type of effective attraction^{15,16}. However, the resulting forces are too small to cause a significant attraction for colloid particles.

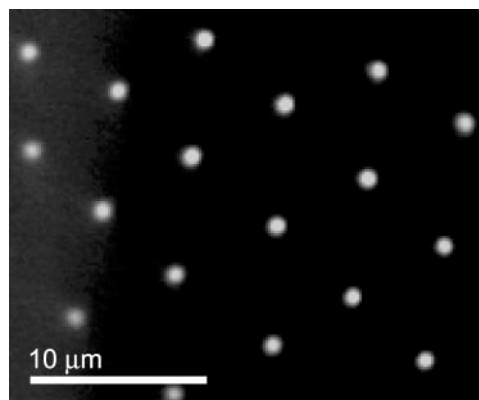


Figure 1 Interfacial colloidal particle ordering induced by repulsive interactions. The figure shows a fluorescence microscopy image of an ordered structure of colloidal particles at an oil–water interface, demonstrating the long range of the repulsive interaction. We use poly(methyl methacrylate) (PMMA) particles, sterically stabilized with poly(hydroxystearic acid)²¹. The particles have a radius of $a = 0.75 \mu\text{m}$ and are suspended in decahydronaphthalene (decalin) at a volume fraction of 0.01 %. The particle cores are labelled with fluorescent dye. An emulsion of water drops in oil is produced by gently shaking around 2 μl of deionized water in 1 ml of the decalin–PMMA mixture, whereupon particles adsorb at the oil–water interface. The particles are strongly bound to the interface owing to the surface energies; this determines the contact angle. They are never observed to escape, which implies that the particles are trapped by a barrier of several tens of $k_B T$. They are observed with a charge-coupled device (CCD) camera mounted on an inverted fluorescence microscope, and the images are recorded on videotape. Sequences of interest are digitized and the particle positions are determined with particle tracking routines²². The mean squared displacement measured at short delay times provides a measure of the particle diffusion coefficient, which is intermediate between that for a particle in the water and in the oil, confirming that the particles are at the interface, partially extending into each of the fluids.

[†] Present address: Lehrstuhl für Biophysik - E22, Technische Universität München, Germany (M.G.N. and A.R.B.); Department of Physics, University of Massachusetts, Amherst, Massachusetts 01003-4525, USA (A.D.D.).

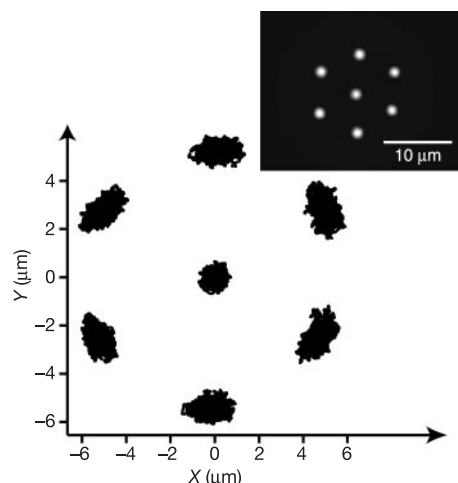


Figure 2 Scatter plot showing positions of a seven-particle hexagonal crystallite on a water droplet of 24 μm radius. The particles maintained this configuration for more than 30 min; 5 min of data, extracted at time intervals of 1/30 s are shown. The inset shows a fluorescent microscope image of the crystallite.

To investigate this question further, we quantified the attractive interaction between like-charged particles at a water–oil interface. A typical configuration of colloidal particles at the interface of a large water drop in oil is shown in Fig. 1. The long range of the repulsive interaction is apparent from the large particle separation. Moreover, this repulsion, combined with the geometric confinement resulting from the finite area of the emulsion drop, which is completely covered with particles, causes the pronounced ordering of the particles. But even when the particle coverage is not complete, ordering can still be observed, as illustrated by the inset in Fig. 2, which shows a group of seven particles in a hexagonal crystallite. These were the only particles on the surface of a large water drop, and the crystallite remained stable for more than 30 min. The persistence of this structure over long times is clear evidence of a long-ranged attractive interaction (Fig. 3).

To extract the interaction potential, we exploit the symmetry of the geometry and measure the probability distribution, $P(r)$, of the centre-to-centre distance, r , of the centre particle to each of the outer particles. To obtain the interparticle potential $V(r)$, we invert the Boltzmann distribution:

$$P(r) \propto \exp \left\{ -\frac{V(r)}{k_B T} \right\} \quad (1)$$

The potential has a minimum at an interparticle separation of $r_{\text{eq}} = 5.7 \mu\text{m}$, and is well described as harmonic with a spring constant of $k = 23 k_B T \mu\text{m}^{-2}$ (Fig. 3).

The motion of any given particle in the crystallite is influenced by the pair interactions with all the other particles; however, by analysing only the radial distance of the outer particles to the centre particle, these many-particle effects are largely cancelled by symmetry. This was confirmed by molecular dynamics simulation of seven particles in the same configuration, interacting with the same pair potential; an uncertainty of about 10–20% in r_{eq} and k was introduced by many-particle effects.

What is the origin of this attraction? A force, F , on the particles normal to the interface distorts the interface, leading to an interparticle capillary attraction:

$$U_{\text{interface}}(r) = \frac{F^2}{2\pi\gamma} \log \left(\frac{r}{r_0} \right) \quad (2)$$

where γ is the interfacial energy and r_0 is an arbitrary constant¹².

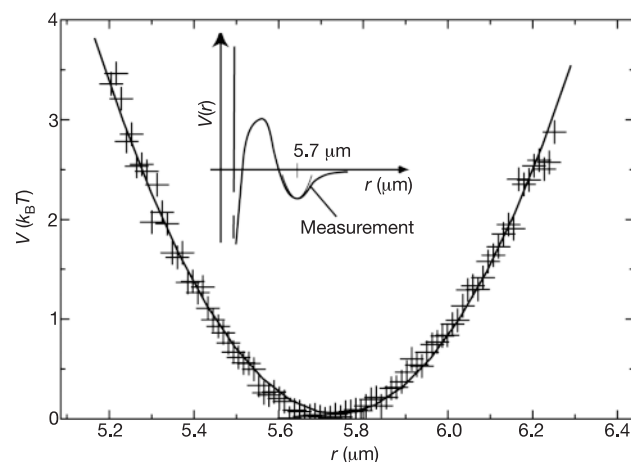


Figure 3 Secondary interparticle potential minimum derived from experimental observations. The data points are obtained from the particle distribution function of the crystallite shown in Fig. 2. The minimum at a separation of 5.7 μm produces the observed hexagonal crystallite. The inset is a sketch of the full interparticle potential and includes the repulsive barrier that stabilizes the particles, and a deep primary minimum at short range that is due to van der Waals attraction. These crystallites were observed to collapse and form a gel after several hours, confirming the presence of the primary minimum and the large repulsive barrier. The accessible range of particle separation allows us to explore the shape of the secondary minimum; the repulsive potential is sufficiently large that we cannot explore the shape or magnitude of the primary attractive well using thermal fluctuations.

Decreasing γ or increasing F increases the interfacial distortion, increasing the capillary attraction. The gravitational force of the colloidal particles on the interface is about 10^{-14} N, far too small to cause the attraction. There is, however, another natural cause for F : The dipolar fields around each particle create electrical stresses that distort the oil–water interface. The sum total of this stress is the net force F that the particle exerts on the interface.

The electrical stresses arise because oil and water have very different dielectric constants: $\epsilon_{\text{oil}} \approx 2$ and $\epsilon_{\text{water}} \approx 80$. When field lines cross the interface, the intensity of the electric field E and the electrostatic energy density $\frac{1}{2}\epsilon\epsilon_0 E^2$ are thus roughly 40 times smaller in water than in oil. Here ϵ_0 is the dielectric constant of vacuum. As a result, the free interface tends to move towards the oil to lower the energy. This is equivalent to a pressure being exerted on the interface towards the oil. The colloidal particle therefore behaves as if pulled into the water by an external force F , as shown schematically in Fig. 4. Thus the particles, which are the source of the electric fields, tend to be surrounded by water, lowering the total electrostatic energy. This effect is analogous to electrostriction¹⁷ and to electro-wetting¹⁸.

The force F can be calculated by the integral of the electric pressure $\frac{1}{2}\epsilon_0\epsilon_{\text{oil}}E_{\text{oil}}^2$ over the free surface. Because the dipolar electric field vanishes like $1/r^3$ and hence the electrostatic pressure vanishes like $1/r^6$, the effect is concentrated very close to the particle, within a distance of order the particle radius a from the contact line. Beyond that distance, the distortion is indistinguishable from any other source of interfacial distortion with the same strength F . A rough estimate of F is obtained by considering that the total electric dipole P of the particle is concentrated at the centre of the water-wetted area. The dipolar field in the oil near the interface is⁹:

$$E_{\text{oil}} \approx \frac{P}{4\pi\epsilon_{\text{oil}}\epsilon_0 r^3} \frac{2\epsilon_{\text{oil}}}{\epsilon_{\text{water}}} \quad (3)$$

The damping factor $2\epsilon_{\text{oil}}/\epsilon_{\text{water}}$ reflects the image charges required for the interface to remain an equipotential. The electric field is not

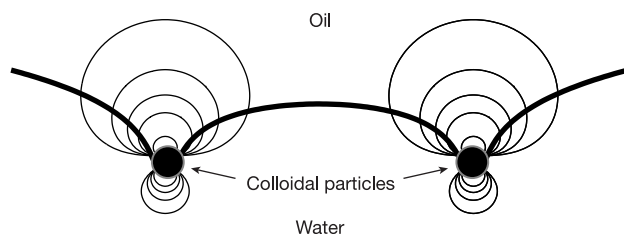


Figure 4 Sketch of the equipotential lines at the fluid interface and the resulting distortion of the oil–water interface. The distortion of the interface shape is greatly enhanced for clarity.

completely screened in the water; the screening ions are part of the charge distribution causing the dipolar fields¹. Integrating the electrostatic pressure from $r = a_w$ to $r = \infty$, where a_w is the radius of the water-wetted region of the particle, we obtain:

$$F \approx \frac{P^2}{16\pi\epsilon_0 a_w^4} \frac{\epsilon_{\text{oil}}}{\epsilon_{\text{water}}^2} \quad (4)$$

$$U = \frac{F^2}{2\pi\gamma} \log\left(\frac{r}{r_0}\right) + \frac{P^2}{4\pi\epsilon_0 r^3} \frac{2\epsilon_{\text{oil}}}{\epsilon_{\text{water}}^2} \quad (5)$$

where the first term in equation (5) is the capillary attraction and the second term is the dipolar repulsion.

From this energy we can predict both r_{eq} and the spring constant, $k = d^2U/dr^2$:

$$\frac{r_{\text{eq}}}{a_w} \approx \left(\gamma a_w^2 \frac{\epsilon_0 a_w^3}{P^2} 768\pi^2 \frac{\epsilon_{\text{water}}^2}{\epsilon_{\text{oil}}} \right)^{1/3} \quad (6)$$

$$\frac{k}{\gamma} \approx \left(\frac{1}{(\gamma a_w^2)^8} \frac{(P^2)^8}{(\epsilon_0 a_w^3)^8} \frac{3}{2^{43}\pi^{13}} \frac{\epsilon_{\text{oil}}^8}{\epsilon_{\text{water}}^{16}} \right)^{1/3} \quad (7)$$

There are two unknown materials parameters in equations (6) and (7). The first is the dipole moment, P , which depends on the screening length, κ^{-1} (taken to be $0.1 \mu\text{m}$), and the surface charge density of the spheres. The second is the radius of the wetted area, a_w , which is determined by the equilibrium contact angle θ , through $a_w = a \sin \theta$, and which depends on the surface charge density.

We obtain agreement between the predicted potential and the experimental data (see Fig. 3) if we adopt a value for a_w that corresponds to an equilibrium contact angle of 26° , and a value for P that corresponds to a total charge on the sphere of 2×10^5 unit charges, or a density σ of 0.6 unit charges per nm^2 . Both of these values are physically reasonable. The fit corresponds to $r_{\text{eq}} = 5.7 \mu\text{m}$ and $k \approx 23 k_B T \mu\text{m}^{-2}$, both in good agreement with the experimentally measured values. The interfacial distortion predicted by the theory is given by the estimate $h \approx F/(2\pi\gamma) \log(r_{\text{eq}}/a)$, and is $h \approx 10 \text{ nm}$, too small to be observed with an optical microscope.

Both the repulsive and the attractive components of the interaction potential depend strongly on the dipole moment, $P = (\sigma/\kappa)\pi a_w^2$. This is directly proportional to the zeta potential of the particle surface, which in turn sensitively depends on surface charge, surface chemistry, and the screening length, κ^{-1} , set by the ion concentration in the water¹⁹. Thus, for example, if the surface charge remains constant, small changes in κ^{-1} change P a great deal, whereas if the surface charge varies and the zeta potential remains constant, there will be no variation of P with κ^{-1} . The attractive interaction is sensitively dependent on P ; thus, if the surface charge remains constant, the attraction becomes negligible compared to $k_B T$ even for low concentrations of added salt, as shown in the Supplementary Information. To verify the sensitivity of the attraction on P , the measurements were repeated with 5 mM NaCl added

to the water; only the repulsive interaction remained and its range was reduced, leading to a correspondingly lower particle stability, as shown in the Supplementary Information. These results also confirm that the particles are charged on the aqueous side, rather than the oil side²⁰. More generally, the existence of a measurable attractive interaction depends on both P and a_w , and thus is sensitive to surface properties, charge and pH, as well as salt concentration.

Although our theory does not fully account for the geometry of the wetted region, it nevertheless captures the essential physics: Dipolar electric fields induce surface charges that distort the interface; the dipolar interaction causes repulsion, while the interfacial distortion causes capillary attraction. This behaviour has broader implications for interfacial and colloid chemistry, because adsorption of charged particles at fluid interfaces is a common phenomenon in foods, drugs, oil recovery, and even biology. $\square$

Received 18 April; accepted 10 September 2002; doi:10.1038/nature01113.

- Pieranski, P. Two-dimensional interfacial colloidal crystals. *Phys. Rev. Lett.* **45**, 569–572 (1980).
- Onada, G. Y. Direct observation of two-dimensional, dynamical clustering and ordering with colloids. *Phys. Rev. Lett.* **55**, 226–229 (1985).
- Denkov, N. D., Veleev, O. D., Kralchevsky, P. A. & Ivanov, I. B. Two-dimensional crystallization. *Nature* **361**, 26 (1993).
- Wickmann, H. H. & Korley, J. N. Colloid crystal self-organization and dynamics at the air/water interface. *Nature* **393**, 445–447 (1998).
- Ruiz-García, J., Gámez-Corralles, R. & Ivlev, B. I. Formation of two-dimensional colloidal voids, soap froths, and clusters. *Phys. Rev. E* **58**, 660–663 (1998).
- Quesada-Pérez, M., Moncho-Jordá, A., Martínez-López, F. & Hidalgo-Álvarez, R. Probing interaction forces in colloidal monolayers: Inversion of structural data. *J. Chem. Phys.* **115**, 10897–10902 (2001).
- Kralchevsky, P. A. & Denkov, N. D. Capillary forces and structuring in layers of colloid particles. *Curr. Opin. Colloid Interf. Sci.* **6**, 383–401 (2001).
- Stamou, D., Duschl, C. & Johannsmann, D. Long-range attraction between colloidal spheres at the air-water interface: The consequence of an irregular meniscus. *Phys. Rev. E* **62**, 5263–5272 (2000).
- Hurd, A. J. The electrostatic interaction between interfacial colloidal particles. *J. Phys. A* **18**, L1055–L1060 (1985).
- Chan, D. Y. C., Henry, J. D. & White, L. R. The interaction of colloidal particles collected at fluid interfaces. *J. Colloid Interf. Sci.* **79**, 410–418 (1981).
- Kralchevsky, P. A., Paunov, V. N., Ivanov, I. B. & Nagayama, K. Capillary meniscus interaction between colloidal particles attached to a liquid–fluid interface. *J. Colloid Interf. Sci.* **151**, 79–94 (1992).
- Morse, D. C. & Witten, T. A. Droplet elasticity in weakly compressed emulsions. *Europhys. Lett.* **22**, 549–555 (1993).
- Bowden, N., Terfort, A., Carbeck, J. & Whitesides, G. M. Self-assembly of mesoscale objects into ordered two-dimensional arrays. *Science* **276**, 233–235 (1997).
- Kralchevsky, P. A. & Nagayama, K. Capillary interactions between particles bound to interfaces, liquid films and biomembranes. *Adv. Colloid Interf. Sci.* **85**, 145–192 (2000).
- Goulian, M., Bruinsma, R. & Pincus, P. Long-range forces in heterogeneous fluid membranes. *Europhys. Lett.* **22**, 145–150 (1993).
- Golestanian, R., Goulian, M. & Kardar, M. Fluctuation-induced interactions between rods on a membrane. *Phys. Rev. E* **54**, 6725–6734 (1996).
- Landau, L. D. & Lifshitz, E. M. *Electrodynamics of Continuous Media* (Addison Wesley Publishing, Pergamon Press, Reading, MA, 1964).
- Berge, B. Electrocapillarity and wetting of insulator films by water. *C. R. Acad. Sci. II* **317**, 2, 157–163 (1993).
- Russel, W. B., Saville, D. A. & Schowalter, W. *Colloidal Dispersions* (Cambridge Univ. Press, Cambridge, UK, 1989).
- Aveyard, R. *et al.* Measurement of long-range repulsive forces between charged particles at an oil–water interface. *Phys. Rev. Lett.* **88**, 246102–1–4 (2002).
- Pusey, P. N. & van Megen, W. Phase-behaviour of concentrated suspensions of nearly hard colloidal spheres. *Nature* **320**, 340–342 (1986).
- Crocker, J. C. & Grier, D. G. Methods of digital microscopy for colloidal studies. *J. Colloid Interf. Sci.* **179**, 298–310 (1996).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements We thank B. Berge and T. M. Squires for discussions. We gratefully acknowledge support from the NSF, the Materials Research Science and Engineering Center through the auspices of the NSF and the Division of Mathematical Sciences. A.B. acknowledges the support from the Emmy Noether-Program of the DFG.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be sent to D.A.W. (e-mail: weitz@deas.harvard.edu).

Interaction of hydrogen with metal nitrides and imides

Ping Chen, Zhitao Xiong, Jizhong Luo, Jianyi Lin & Kuang Lee Tan

Physics Department, National University of Singapore, 10 Kent Ridge Crescent, 119260 Singapore

The pursuit of a clean and healthy environment has stimulated much effort in the development of technologies for the utilization of hydrogen-based energy. A critical issue is the need for practical systems for hydrogen storage, a problem that remains unresolved after several decades of exploration. In this context, the possibility of storing hydrogen in advanced carbon materials has generated considerable interest. But confirmation and a mechanistic understanding of the hydrogen-storage capabilities of these materials still require much work^{1–5}. Our previously published work on hydrogen uptake by alkali-doped carbon nanotubes cannot be reproduced by others^{6–8}. It was realized by us and also demonstrated by Pinkerton *et al.*⁸ that most of the weight gain was due to moisture, which the alkali oxide picked up from the atmosphere. Here we describe a different material system, lithium nitride, which shows potential as a hydrogen storage medium. Lithium nitride is usually employed as an electrode, or as a starting material for the synthesis of binary or ternary nitrides^{9,10}. Using a variety of techniques, we demonstrate that this compound can also reversibly take up large amounts of hydrogen. Although the temperature required to release the hydrogen at usable pressures is too high for practical application of the present material, we suggest that more investigations are needed, as the metal–N–H system could prove to be a promising route to reversible hydrogen storage.

Temperature dependences of hydrogen absorption and desorption in the fresh Li_3N sample were investigated by thermogravimetry (Fig. 1). It can be seen that the absorption of hydrogen begins at a temperature of around 100 °C. In the temperature range 170–210 °C, there is a fast weight gain. A total increase of about 9.3 wt% was achieved after maintaining the sample at 255 °C for half an hour. We noted that if the temperature was kept long enough below 200 °C, the Li_3N sample could also absorb a substantial amount of hydrogen. The hydrogenated sample was cooled down to 50 °C and evacuated to 10^{-5} mbar. The desorption process appeared to be composed of two parts: under high vacuum, a large portion (6.3 wt%) of hydrogen was released at temperatures below 200 °C; the remaining 3 wt% hydrogen could be desorbed only when the temperature was raised above 320 °C. Mass spec-

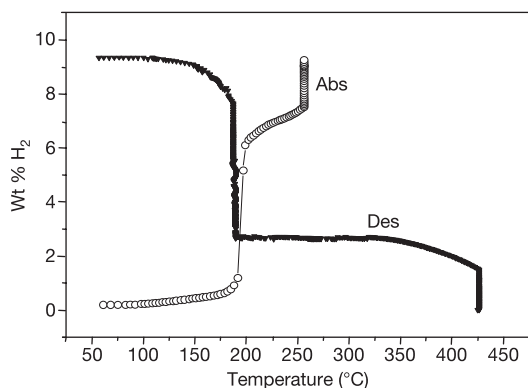


Figure 1 Weight variations during hydrogen absorption and desorption processes over Li_3N samples, details are given in the Methods. Abs, absorption; Des, desorption.

trometry and an ammonia-sensitive reagent were applied to identify the released gas. Only hydrogen was found to be released in the tested temperature range (see Supplementary Information).

Pressure–composition (P – C) isotherms of fresh Li_3N samples were measured at several temperatures. Those recorded at 195 °C, 230 °C and 255 °C are presented in Fig. 2. As illustrated in the absorption isotherms, about 3.6 hydrogen atoms can be absorbed by one Li_3N molecule, that is, about 10 wt% of the sample. The hydrogenation degree can be further improved if the starting material Li_3N is free of contamination. Unlike most metal hydrides, which present one plateau in the P – C isotherm, Li_3N has two. The first one (in the range $0 < \text{H}/\text{Li}_3\text{N} < 1.1$) has rather low equilibrium pressure (below 0.07 bar), so we deduce that hydrogen absorbed in this region might not be easily desorbed, and should correspond to the high-temperature-desorbed portion detected by thermogravimetry. The second plateau is sloped, the overall equilibrium pressure is below 0.2 bar at 195 °C, 0.5 bar at 230 °C and 1.5 bar at 255 °C. The desorption isotherms, as predicted, cannot return to the origin. Under P – C isotherm conditions (desorption stopped at 0.04 bar), about 55% hydrogen can be desorbed at temperatures above 230 °C, which is less than that detected by thermogravimetry under high vacuum (10^{-5} mbar). We noticed that after the stop of

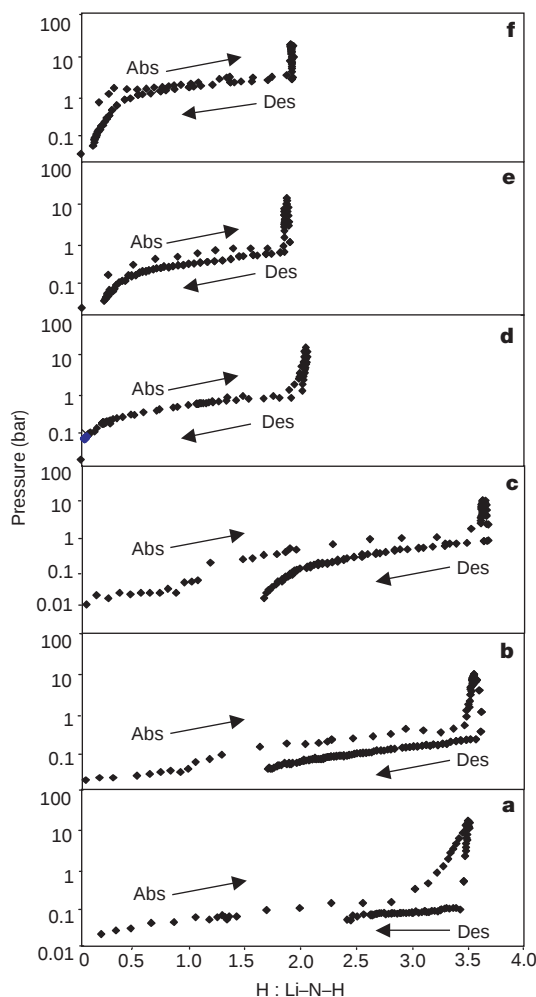
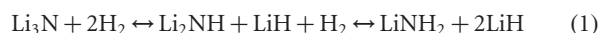


Figure 2 Pressure–composition (P – C) isotherms of Li_3N and Li_2NH samples. Pressure was increased step by step to 20 bar then gradually reduced to 0.04 bar, other details are given in the Methods. The x axis represents the molar ratio of H atom to Li–N–H molecule. **a**, Li_3N at 195 °C; **b**, Li_3N at 230 °C; **c**, Li_3N at 255 °C; **d**, Li_3N re-PCI at 255 °C; **e**, Li_2NH at 255 °C and **f**, Li_2NH at 285 °C.

P–*C* isotherm measurements, there was continuous hydrogen desorption from the remaining material under lower pressures, indicating that hydrogen desorption did not complete at that stage. Hysteresis in desorption could be observed; however, after a few cycles of PCIs, it was found to be much improved. As can be seen in Fig. 2d, the repeated absorption–desorption isotherms of the Li_3N sample almost coincide.

To facilitate the understanding of phase and composition changes during hydrogen absorption and desorption in the Li_3N sample, we collected pristine, half-hydrogenated, fully-hydrogenated, half-dehydrogenated and fully-dehydrogenated Li_3N samples for X-ray diffraction measurements. Results (Fig. 3) show that the pristine hexagonal Li_3N phase shifted to face-centred cubic (f.c.c.) lithium imide and hydride phases¹¹ after being half-hydrogenated. The fully-hydrogenated Li_3N sample is composed of body-centred tetragonal lithium amide¹¹, that is, LiNH_2 and the enhanced lithium hydride phase. Phase changes in desorption follow a process almost the reverse of that of absorption. Thus, we deduce that the reversible hydrogen storage in Li_3N sample may take the following reaction path:



The molar ratio of absorbed hydrogen atoms to Li_3N molecule is four: this works out to be 11.5 wt% of Li_3N , corresponding well to that obtained by *P*–*C* isothermal and thermogravimetric measurements. The values of standard enthalpy of formation for Li_3N , Li_2NH , LiNH_2 and LiH from the elements are -197 kJ mol^{-1} , -222 kJ mol^{-1} , -176 kJ mol^{-1} and -91 kJ mol^{-1} (ref. 12), respectively. The overall reaction heat is -161 kJ mol^{-1} .

In 1910, Dafert and Miklausz reported that, when Li_3N reacted with H_2 , Li_3NH_4 was formed¹³:



In fact, Li_3NH_4 is the mixture of LiNH_2 and 2LiH . No investigation on the reverse reaction was done. The reason may lie in the known fact that lithium amide will decompose to lithium imide and ammonia at increased temperatures^{14,15}. The existence of LiH seems anomalous (see Supplementary Information).

Inferred from *P*–*C* isothermal and X-ray diffraction results, Li_2NH could reversibly store more hydrogen at temperatures below 300°C . Figures 2e and f are the *P*–*C* isotherms of the

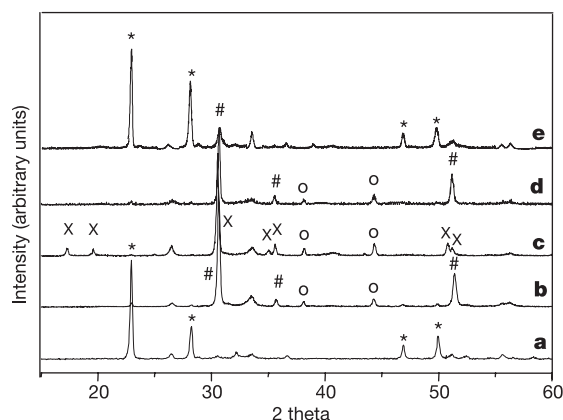


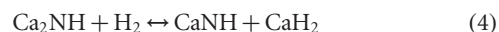
Figure 3 Structure changes during hydrogen absorption and desorption processes. **a**, Pristine Li_3N ; **b**, half-hydrogenated sample; **c**, fully hydrogenated sample; **d**, half-dehydrogenated sample and **e**, completely dehydrogenated sample. Peaks marked with an asterisk are the characteristic diffraction of Li_3N , crosses represent LiNH_2 , hashes represent Li_2NH and circles represent LiH . Small features around $32\text{--}34^\circ$ and $55\text{--}56^\circ$ are the diffraction peaks of LiOH and Li_2O ; the feature at around 26° is carbon. Sample collection and details for measurements are described in the Methods.

Li_2NH sample recorded at 255°C and 285°C , respectively. About 1.85 hydrogen atoms, which is equal to about 6.5 wt% of sample, was stored. Almost all hydrogen absorbed can be desorbed. Hydrogen storage leads to the formation of LiNH_2 and LiH , so:



We see that, theoretically, ~ 7.0 wt% of hydrogen can be reversibly stored in Li_2NH .

A similar hydrogen storage phenomenon was observed in the Ca – N – H system. It was demonstrated that Ca_2NH could reversibly store hydrogen in the temperature range $350\text{--}600^\circ\text{C}$ (see Supplementary Information). *P*–*C* isotherms measured at 500°C and 550°C , respectively, are presented in Fig. 4. About 1.8 hydrogen atoms, which is equal to ~ 1.9 wt% of sample, can be reversibly stored in Ca_2NH . The hydrogenated Ca_2NH sample is composed of CaNH and CaH_2 , indicating that hydrogen is stored following the reaction:



Thus, theoretically, about 2.1 wt% of hydrogen can be reversibly stored in Ca_2NH .

The general features of isotherms for Li_2NH and the recyclable portion of Li_3N at temperatures below 300°C resemble each other closely, indicating similar thermodynamic properties. Thermodynamic analysis was performed over the Li_2NH and the Ca_2NH systems. As all the absorption/desorption isotherms are slightly sloped, we chose the middle points for calculation. Using van't Hoff plots (see Supplementary Information), it is revealed that the standard enthalpy of absorption (ΔH_{ab}) is about $-66.1 \text{ kJ mol}^{-1}$ for Li_2NH and $-88.7 \text{ kJ mol}^{-1}$ for Ca_2NH . In Li – N – H systems, the absorption and desorption isotherms coincide after a few cycles, so the calculated ΔH_{abs} should be equal to ΔH_{des} . To raise the plateau pressure up to 1.0 bar, a temperature above 275°C is required. This temperature is on the high side for the immediate applications of the system. The hydrogenation temperature ranges for Li_2NH , Li_3N and Mg_2NiH_4 ($\Delta H = -64.4 \text{ kJ mol}^{-1}$) are close to each other¹⁶, while Li_3N and Li_2NH possess much higher storage capacity than Mg_2Ni (~ 11.5 wt% and 7.0 wt% versus ~ 3.8 wt%).

Li_3N , Li_2NH and Ca_2NH have fast kinetics in hydrogen storage. For a 500 mg sample, almost all hydrogen (for Li_2NH and Ca_2NH) or a substantial amount of hydrogen (for Li_3N) can be absorbed within 10 min under 30 bar of hydrogen and at 255°C (for Li_3N and Li_2NH) and 500°C (Ca_2NH), respectively (see Supplementary Information). Desorption rate strongly depends on temperature and hydrogen pressure. All these three systems can be recycled: after

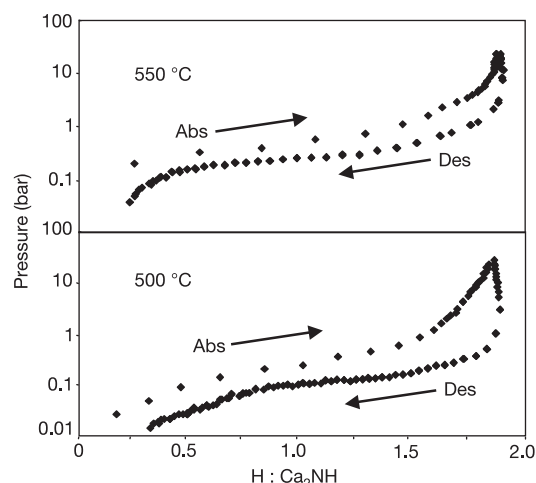


Figure 4 *P*–*C* isotherms of the Ca_2NH sample measured at 500°C and 550°C , respectively. For details of measurement, please refer to the Methods.

cycles of absorption/desorption, the overall capacity fluctuates within 10% (see Supplementary Information).

Elements, especially those in groups I–IV and some transition metals, have their nitride, hydride and amide/imide forms. There is, therefore, still plenty of scope for further exploring metal–N–H systems for hydrogen storage. In the present work, we have found that Li–N–H and Ca–N–H systems provide interesting methods of reversible hydrogen storage. However, to meet practical applications at more moderate temperatures and with improved chemical stability, further work is needed for better material design, engineering and mechanistic understanding. □

Methods

Lithium nitride was synthesized by heating commercial lithium metal under a purified nitrogen atmosphere at 100 °C overnight. Owing to the contamination introduced in loading and transferring of the sample, the degree of nitrogenation is less than 95%. The Li₂NH sample was synthesized by decomposing commercial LiNH₂ (95% purity) at temperatures above 350 °C overnight. The Ca₂NH sample was prepared by hydrogenation followed by dehydrogenation of Ca₃N₂ + CaH₂ (1:1) mixture at temperature around 550 °C.

P–C Isotherm measurements

Hydrogen uptake by Li–N–H and Ca–N–H systems was measured with a commercial pressure–composition isotherm (PCI) unit provided by Advanced Materials (<http://www.advanced-material.com>). Static P–C isotherms were determined at temperatures of 195 °C, 230 °C, 255 °C, 285 °C for Li–N–H samples and 500 °C and 550 °C for Ca₂NH samples. The delay time was extended to 200 s, that is, data was recorded only when there was no pressure change within 200 s. The dimensions of the sample cell are about 1.0 cm in diameter and about 3.0 cm in height. A sample of approximately 500 mg was tested each time. A tube furnace with accuracy of 0.1 °C was used as the heater. A thermocouple was closely contacted to the outer surface of the hydrogenation chamber. The possible uncertainties of P–C isothermal measurements may originate from the following three aspects: (1) a temperature difference between sample and thermocouple, estimated to be in the range of ±5.0 °C; (2) the variation in sample density during absorption and desorption; and (3) contamination introduced in the sample loading and gaseous phase.

Thermogravimetric measurements

Weight variation within hydrogenation and dehydrogenation processes was monitored with an intelligent gravimetric analyzer (IGA) from Hiden (<http://www.hiden.co.uk>). Temperature was measured with a thermocouple placed slightly above the sample. A sample of about 200 mg was tested. In the hydrogenation process, about 3.0 bar of purified hydrogen was introduced into sample chamber; the temperature was gradually increased from 50 °C to 255 °C at 2 °C min^{−1} intervals and was maintained at 255 °C for 30 min. In the dehydrogenation process, the hydrogenated sample was cooled to 50 °C. The sample chamber was evacuated to 10^{−5} mbar, and the sample temperature was gradually increased to 195 °C and kept constant until weight loss was almost undetectable. After this, sample temperature was further increased to 430 °C. The uncertainties in thermogravimetric measurements mainly come from the contamination of the sample during loading when exposure to air is unavoidable. The temperature difference between sample and thermocouple was around ±5.0 °C.

Desorption and structural measurements

Temperature-programmed desorption (TPD, with purified Ar as carrier gas) was conducted on a home-made Reactor–MS–GC combined system. Around 300 mg of sample was tested each time. Mass spectrometry (MS) and gas chromatography (GC) were applied to detect the outlet gases. Because ammonia is probably formed during desorption, the outlet gas was conducted into Nessler's reagent to identify any trace of NH₃. A Bruker D8-advance X-ray diffractometer with CuKα radiation was used to identify structural/compositional changes. Except for the complete dehydrogenated Li₃N sample, which was collected by heating the pre-hydrogenated sample to 430 °C under high vacuum, other samples at different degrees of hydrogenation and dehydrogenation were obtained after various stages of PCI measurements (at 255 °C).

Received 7 June; accepted 10 October 2002; doi:10.1038/nature01210.

1. Dillon, A. C. *et al.* Storage of hydrogen in single-walled carbon nanotubes. *Nature* **386**, 377–379 (1997).
2. Liu, C. *et al.* Hydrogen storage in single-walled carbon nanotubes at room temperature. *Science* **286**, 1127–1129 (1999).
3. Ye, Y. *et al.* Hydrogen adsorption and cohesive energy of single-walled carbon nanotubes. *Appl. Phys. Lett.* **74**, 2307–2309 (1999).
4. Hirscher, M. *et al.* Hydrogen storage in sonicated carbon materials. *Appl. Phys. A* **72**, 129–132 (2001).
5. Mereghetti, V. & Parrinello, M. Review of theoretical calculations of hydrogen storage in carbon-based materials. *Appl. Phys. A* **72**, 143–146 (2001).
6. Chen, P., Wu, X., Lin, J. & Tan, K. L. High H₂ uptake by alkali-doped carbon nanotubes under ambient pressure and moderate temperature. *Science* **285**, 91–93 (1999).
7. Yang, R. T. Hydrogen storage by alkali-doped carbon nanotubes-revisited. *Carbon* **38**, 623–626 (2000).
8. Pinkerton, F. E. *et al.* Thermogravimetric measurement of hydrogen absorption in alkali-modified carbon materials. *J. Phys. Chem. B* **104**, 9460–9467 (2000).

9. O'Loughlin, J. L., Wallace, C. H., Knox, M. S. & Kaner, R. B. Rapid solid-state synthesis of tantalum, chromium, and molybdenum nitrides. *Inorg. Chem.* **40**, 2240–2245 (2001).
10. Shodai, T., Okada, S., Tobishima, S. & Yamai, J. Anode performance of a new layered nitride Li_{3-x}Co_xN (x = 0.2–0.6). *J. Power Source* **68**, 515–518 (1997).
11. *Power Diffraction File TM Data sets: 1–49* (International Center for Diffraction Data (ICDD), Pennsylvania, USA, 1999).
12. *Lithium, Gmelins Handbuch-Der Anorganischen Chemie* System Number 20 (ed. Meyer, R. J.) 273–270 (Verlag Chemie, GMBH, Weinheim/Bergstrasse, 1960).
13. Däfer, F. W. & Miklausz, R. Über einige neue verbindungen von stickstoff und wasserstoff mit lithium. *Monatsch. Chem.* **31**, 981–996 (1910).
14. Juza, R. & Opp, K. Metallic amides and metallic nitrides. XXIV. The crystal structure of lithium amide. *Z. Anorg. Allg. Chem.* **266**, 313–324 (1951).
15. Schenk, P. W. *Nitrogen, Handbook of Preparative Inorganic Chemistry* 464–465 (Academic Press, New York, 1963).
16. Jung, W. B., Nahm, K. S. & Lee, W. Y. The reaction-kinetics of hydrogen storage in Mg₂Ni. *Int. J. Hydrogen Energy* **15**, 641–648 (1990).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com/nature>).

Acknowledgements We thank A. Nazri and the General Motors R&D Centre (Warren, Detroit, USA) for the facilitation of confirmation tests. The work is financially supported by the Agency for Science, Technology and Research (A*STAR) of Singapore.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.C. (e-mail: phychenp@nus.edu.sg).

Evidence for recycled Archaean oceanic mantle lithosphere in the Azores plume

Bruce F. Schaefer*, Simon Turner†, Ian Parkinson*, Nick Rogers* & Chris Hawkesworth†

* Department of Earth Sciences, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK

† Department of Earth Sciences, Wills Memorial Building, University of Bristol, Bristol BS8 1RJ, UK

The compositional differences between mid-ocean-ridge and ocean-island basalts place important constraints on the form of mantle convection^{1,2}. Also, it is thought that the scale and nature of heterogeneities within plumes and the degree to which heterogeneous material endures within the mantle might be reflected in spatial variations of basalt composition observed at the Earth's surface. Here we report osmium isotope data on lavas from a transect across the Azores archipelago which vary in a symmetrical pattern across what is thought to be a mantle plume. Many of the lavas from the centre of the plume have lower ¹⁸⁷Os/¹⁸⁸Os ratios than most ocean-island basalts and some extend to subchondritic ¹⁸⁷Os/¹⁸⁸Os ratios—lower than any yet reported from ocean-island basalts. These low ratios require derivation from a depleted, harzburgitic mantle, consistent with the low-iron signature of the Azores plume. Rhenium-depletion model ages extend to 2.5 Gyr, and we infer that the osmium isotope signature is unlikely to be derived from Iberian subcontinental lithospheric mantle. Instead, we interpret the osmium isotope signature as having a deep origin and infer that it may be recycled, Archaean oceanic mantle lithosphere that has delaminated from its overlying oceanic crust. If correct, our data provide evidence for deep mantle subduction and storage of oceanic mantle lithosphere during the Archaean era.

Mantle plumes are major sources of magmatism, and they provide windows into the deep mantle. Elevated $^{206}\text{Pb}/^{204}\text{Pb}$ and low $^{143}\text{Nd}/^{144}\text{Nd}$ isotope ratios of some ocean-island basalts provide evidence for recycling of ancient, enriched crustal materials³, and much work has been done on Hawaii, which is a spectacular product of plume-related magmatism. However, Hawaii is unusual because it is situated above a plume of high-buoyancy flux, and at the present time most plumes are very much weaker⁴. In the Atlantic, Iceland represents a significant volume of magmatism astride the Mid-Atlantic Ridge, and the Azores islands make for an interesting contrast because although they also lie above a plume that intersects the Mid-Atlantic Ridge, the islands are scattered either side of the ridge rather than on it.

Idealized plumes are axisymmetric, and White *et al.*⁵ demonstrated a concentric isotope pattern in the Galapagos that was attributed to the enfolding of surrounding mid-ocean-ridge (MORB)-type mantle by the ascending plume. This in turn provided an important link to fluid dynamic models of plume emplacement. However, partial melts from lower-buoyancy plumes, where the amounts and rates of melting are much less, are more likely to preserve information on the nature and scale of the chemical variations and the local causes of melting within plumes. Thus, it is intriguing that the areas of thickest crust in the Azores are neither astride the ridge, where the extension and hence the amounts of melting should be greatest, nor are the islands distributed symmetrically about the ridge (Fig. 1a). One possible

explanation is that the extent and rates of melting are not symmetrical owing to the presence of localized, more volatile source regions beneath the individual islands, and in particular below São Miguel^{6,7}. Previous work on the Azores has confirmed the presence of compositional heterogeneities^{7–12}, and three source components have been identified; MORB-source mantle, a regional plume component and an enriched component most strongly evident in the lavas on São Miguel. This last component has been inferred either to contain a contribution from recycled sediment^{7,10}, or to represent localized pieces of recycled, hydrous, subcontinental, lithospheric mantle entrained within the plume¹². The former interpretation implies a relatively homogeneous plume interacting with MORB-mantle and sediment in the shallow upper mantle, whereas the latter involves heterogeneities that remain spatially and compositionally distinct and asymmetrically distributed in the rising plume. Here we present an Os isotope transect across the Azores plume that explores further the nature and origin of components sampled in this distinctive mantle plume.

The new Os isotope data are from well-characterized, historic samples from the Azores^{7,11}, and are presented in Table 1 along with details of the analytical techniques used. Previous Os studies¹² in the Azores have focused largely on São Miguel, and our data from that island similarly reveal elevated $^{187}\text{Os}/^{188}\text{Os}$ ratios (0.133–0.142) indicative of the involvement of an enriched component. Away from São Miguel, some lavas have $^{187}\text{Os}/^{188}\text{Os}$ ratios in the range 0.126–0.139, which is similar to that observed in MORB^{13,14}. However, the striking result is that, excluding São Miguel, 50% of the Azores lavas have $^{187}\text{Os}/^{188}\text{Os} < 0.127$ whereas only 3 of all the data previously published from ocean island basalts¹⁴ have ratios below this value. Moreover, some lavas from Pico and Faial, in the centre of the Azores, extend down to $^{187}\text{Os}/^{188}\text{Os} = 0.110$. These represent the lowest $^{187}\text{Os}/^{188}\text{Os}$ ratios yet measured in oceanic lavas¹⁴. The lithospheric mantle beneath the present-day oceanic plates may, over time, develop subchondritic Os isotope ratios. However, the lithosphere through which the Azores lavas were erupted is far too young (<15 Myr) to have developed the extreme subchondritic Os isotope ratios that are observed, and so these ratios are unlikely to reflect assimilation.

^{187}Re is incompatible relative to its daughter isotope (^{187}Os) during mantle melting: in order for a mantle source region to develop strongly subchondritic Os isotope ratios, ~25% melt extraction is required—so that most of the Re is removed—followed by ageing for about 1 Gyr or more^{14–16}. This degree of melt extraction will effectively exhaust a peridotite source of clinopyroxene, leaving behind an Fe-poor, harzburgitic residue. This is consistent with the low-Fe signature of the Azores plume^{7,8} and positive ϵ_{Nd}

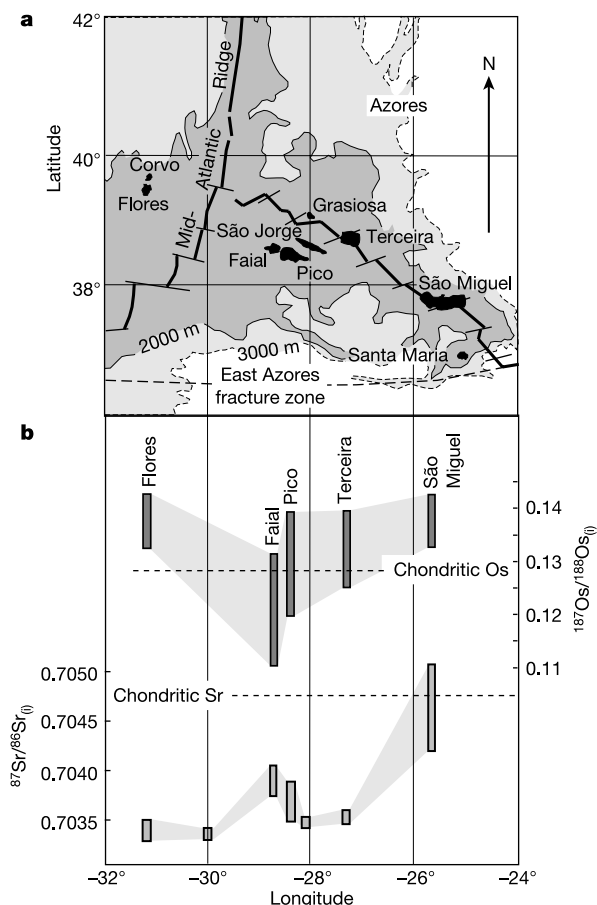


Figure 1 Map of the Azores archipelago showing bathymetry and islands sampled (a), along with the average $^{187}\text{Os}/^{188}\text{Os}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratio for each island plotted along a northwest-southeast transect across the plume (b). Data from this study and compiled from ref. 12. (i), initial.

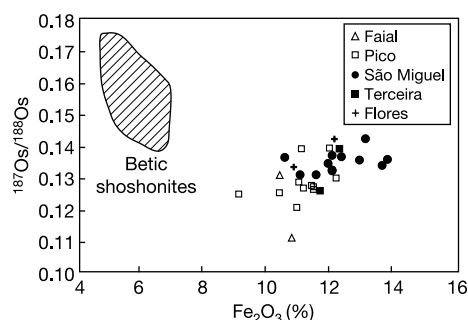


Figure 2 Fe_2O_3 content versus $^{187}\text{Os}/^{188}\text{Os}$ ratio, showing the broad positive correlation ($r^2 = 0.3$) in the Azores lavas. This correlation is consistent with a long-term, time-integrated, depleted nature of the Azores plume component. Field for the Betic shoshonites¹⁶ indicates that the Iberian sub-continental lithospheric mantle cannot be the origin of this component.

Table 1 Re-Os data for the Azores

	MgO (%)	[Ni] (p.p.m.)	[Yb] (p.p.m.)	[Re] (p.p.b.)	[Os] (p.p.b.)	¹⁸⁷ Os/ ¹⁸⁸ Os	2σ _m *	¹⁸⁷ Re/ ¹⁸⁸ Os	2σ _m †
Sao Miguel									
S1	7.76	89	2.314	0.347	0.0252	0.13528	0.00037	66.45	0.07
S3	8.34	154	2.350	0.359	0.0797	0.13517	0.00083	21.70	0.32
S10	8.33	100	2.055	0.380	0.0173	0.14196	0.00026	106.11	0.11
SP2	7.93	96	2.279	0.723	0.0226	0.13338	0.00043	154.21	0.15
SP8	9.10	160	2.173	0.271	0.0580	0.13413	0.00026	22.50	0.02
Terceira									
T6	8.70	122	2.269	0.261	0.0107	0.13875	0.00027	117.64	0.12
T18	7.94	121	2.278	0.253	0.1476	0.12558	0.00009	8.23	0.01
Pico									
P4	10.36	206	1.810	0.109	0.0228	0.13863	0.00023	22.98	0.02
P5	9.63	165	1.941	0.043	0.0369	0.12012	0.00017	5.58	0.01
P25	8.24	119	2.024	0.220	0.0127	0.12631	0.00029	83.17	0.08
P26	10.83	195	1.876	0.077	0.0197	0.12670	0.00020	18.77	0.02
P29	8.16	131	2.309	0.179	0.0350	0.12493	0.00015	24.63	1.41
P2	18.32	398			0.0483	0.12445	0.00017		
P7	10.12	167			0.0229	0.12584	0.00036		
P10	10.20	154	2.23		0.0045	0.13885	0.00042		
P14	9.03	134			0.0076	0.12957	0.00029		
P23	7.99	100	2.37		0.0095	0.12811	0.00030		
P27	10.87	193			0.0297	0.12709	0.00045		
Faial									
F/CP-18	7.83	114	1.979	0.042	0.0125	0.13058	0.00025	16.41	3.70
F/CA-6	9.73	170	1.812	0.150	0.0164	0.11019	0.00039	43.75	0.04
Flores									
FL8	8.45	128	1.807		0.0718	0.14180	0.00197		
FL20	9.17	134	1.711		0.0362	0.13298	0.00042		

Samples were spiked with ¹⁹⁰Os and ¹⁸⁵Re tracers followed by Carius tube dissolution²¹, solvent extraction²², and analysis on a Finnigan MAT 261 mass spectrometer in negative thermal ionization mass spectrometry (N-TIMS) mode at the Open University. Long-term ¹⁸⁷Os/¹⁸⁸Os reproducibility ≤0.2% (2 s.d.) as determined from routine solution standard measurement (solution Os (Department of Terrestrial Magnetism, Washington, USA (DTM)); ¹⁸⁷Os/¹⁸⁸Os = 0.173963 ± 0.000171 (2σ); n = 140 from 10 Feb. 1997 to 19 Oct. 2000); ¹⁸⁷Os/¹⁸⁸Os normalized to ¹⁹²Os/¹⁸⁸Os = 3.08262. All data are blank corrected; Os total procedural blank (TPB) levels were <1 pg, resulting in blank contributions <1% and consistently <0.5%. Particularly low abundance samples had blank corrections of up to 1.5% on ¹⁸⁷Os/¹⁸⁸Os. TPBs were run as every ninth sample during the course of this study, with ¹⁸⁷Os/¹⁸⁸Os = 0.1683 (n = 3), allowing ¹⁸⁷Os/¹⁸⁸Os correction with a high level of confidence. Blank corrections for Re were more variable, and were typically 2–10%, with two samples ranging up to 25%. As TPBs were run as part of each batch of dissolutions, the appropriate blank correction for each batch was applied.

*2σ error cited from within run precision.

†¹⁸⁷Re/¹⁸⁸Os uncertainties propagated from within run counting statistics to a minimum of 1%.

values^{7,10,11}, which independently argue for a melt-depleted source. Figure 2 illustrates that there is also a broad positive correlation between Fe content and ¹⁸⁷Os/¹⁸⁸Os ratios in the lavas analysed. Re-depletion model ages can be used to estimate the timing of depletion¹⁴, and these extend to 1–2.5 Gyr ago for the two lowest ¹⁸⁷Os/¹⁸⁸Os lavas from Pico and Faial (Fig. 3). Combining the data from all of the islands, the range of Os isotopes in the Azores is extremely large.

Whereas the subchondritic Os isotope ratios and positive ε_{Nd} values reflect long-term depletion of a component in the Azores plume, all of the lavas are also enriched in incompatible elements; such enrichment requires a relatively recent metasomatic event^{7,10,11} that must have also added volatiles in order to facilitate later melting of the depleted peridotite¹⁷. Depleted-mantle Nd model ages suggest that this event may have occurred at around 300–550 Myr ago⁷, and because such metasomatism would also have added Re, and possibly Os, it may have been responsible for some of the large range in Os isotopes. Given this evidence for a complex history of depletion and subsequent metasomatism, good correlations between Os isotopes and other isotope or incompatible-trace-element ratios are not to be expected, and the symmetry observed in the Os isotope data is less clear in other radiogenic isotope ratios (for example, ⁸⁷Sr/⁸⁶Sr, Fig. 1b). However, this also reflects the sampling of an enriched, possibly sedimentary^{7,10} component, which by its localized and non-symmetric distribution, is inferred to lie in the shallow mantle beneath São Miguel. Because the above-mentioned Re model ages assume evolution with no radiogenic in-growth, and the processes of metasomatism and mixing are only likely to have resulted in subsequent increases in the ¹⁸⁷Os/¹⁸⁸Os ratios, these model ages represent minimum ages for the depleted plume component that must therefore be at least late Archaean in age (Fig. 3). The key question is the origin of this depleted component.

An important finding is that the Os isotope data are symmetrical along a northwest–southeast transect across the Azores bathymetric swell (Fig. 1b). The composition of some Azores lavas may reflect a contribution from the MORB-source⁷, and the more elevated Os isotope ratios at the plume margins are consistent with, but do not demand, this. The subchondritic Os isotope ratios are a feature of the centremost islands, and so this is a characteristic attributed to the deep, central rising plume material. This provides a conservative estimate of ~100 km for the radial distribution of this component within the plume. These data also require derivation from an ancient, strongly depleted, harzburgitic source. It has previously been suggested that the enriched, São Miguel component may be Iberian, subcontinental lithospheric mantle that became smeared-out in the shallow mantle during the opening of the North Atlantic¹². However, the subchondritic Os isotope ratios are found in the centre of the Azores, rather than on São Miguel, and it appears that the subchondritic Os material is an integral part of the plume. As the overall age of the Iberian lithosphere is mid-Proterozoic and the Os age of the Azores low-Fe component is late Archaean, or older, it is unlikely that the mantle lithosphere from the continent adjacent to the site of break-up is the source of the low Os isotope ratios. Moreover, a recent Os isotope study of potassic lavas inferred to be produced by heating during convective removal of enriched, Iberian subcontinental, lithospheric mantle¹⁶ has shown that, although such materials have low Fe content, they have ¹⁸⁷Os/¹⁸⁸Os ratios that are far too high to form a suitable end-member for the Azores array (Fig. 2).

An alternative explanation is that each Azores island represents localized melting of pieces of convectively removed, subcontinental lithospheric mantle entrained from depth within the plume⁶. The combined symmetry of the Os-isotope and low-Fe (refs 7, 8) signatures argues against this, and suggests that the depleted

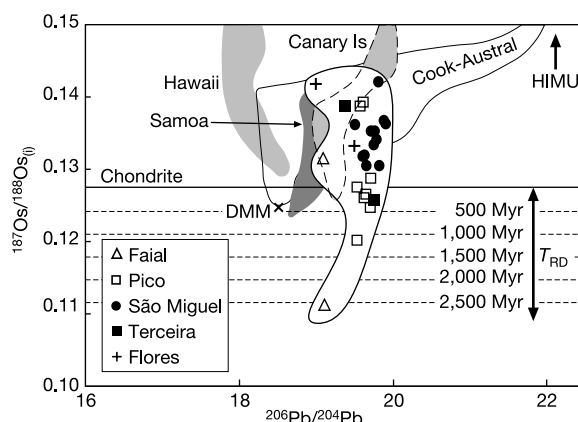


Figure 3 Plot of $^{187}\text{Os}/^{188}\text{Os}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ for the Azores lavas (including data from ref. 12). Also shown are Re-depletion model ages (T_{RD}) and comparison with the fields for other ocean islands^{23–30}. HIMU, high μ ; DMM, depleted MORB mantle.

component is a feature of the regional, deep plume signal⁷ rather than localized beneath particular islands. In principle, this depleted component could be oceanic or continental lithosphere. Subchondritic Os isotope ratios are particularly characteristic of Archaean subcontinental lithospheric mantle xenoliths, and they can demonstrably survive the effects of subsequent metasomatism recorded by elevated Sr and Pb isotopes^{14,15}. However, one of the reasons inferred for the stability and elevation of Archaean cratons is that such lithosphere is rendered highly buoyant by the strong depletion of Fe that results from high degrees of melting during their formation. The corollary is that such materials appear to be very resistant to convective removal¹⁸. Therefore, we are intrigued by the possibility that the subchondritic Os ratios from the centre of the Azores may represent recycled, ancient oceanic mantle lithosphere. Because the mantle was hotter in the Archaean, the oceanic crust would have been thicker and its lithosphere more depleted than at the present day, although not as Fe-poor as Archaean subcontinental lithospheric mantle (as inferred from mantle xenoliths¹⁹). As ancient oceanic crust would have evolved to very high present-day $^{206}\text{Pb}/^{204}\text{Pb}$ ratios, not observed in the Azores (Fig. 3), such oceanic mantle lithosphere would have to have been delaminated from its overlying oceanic crust²⁰.

Thus a straightforward, but not necessarily unique, explanation is that the Azores plume represents sampling of the harzburgitic lithosphere from an Archaean oceanic plate that was subducted into the deep mantle. If correct, this suggests that compositional heterogeneities can have great longevity in the mantle, and provides evidence for deep mantle subduction and storage of oceanic mantle lithosphere since the Archaean. □

Received 20 May; accepted 1 October 2002; doi:10.1038/nature01172.

- Gast, P. W., Tilton, G. R. & Hedge, C. Isotopic composition of lead and strontium from Ascension and Gough Islands. *Science* **145**, 1181–1185 (1964).
- Hofmann, A. W. Mantle geochemistry: the message from oceanic volcanism. *Nature* **385**, 219–229 (1997).
- Zindler, A. & Hart, S. Chemical geodynamics. *Annu. Rev. Earth Planet. Sci. Lett.* **14**, 493–571 (1986).
- Sleep, N. H. Hotspots and mantle plumes: some phenomenology. *J. Geophys. Res.* **95**, 6715–6736 (1990).
- White, W. M., McBirney, A. R. & Duncan, R. A. Petrology and geochemistry of the Galápagos Islands: portrait of a pathological mantle plume. *J. Geophys. Res.* **98**, 19533–19563 (1993).
- McKenzie, D. & O'Nions, R. K. The source regions of ocean island basalts. *J. Petrol.* **36**, 133–159 (1995).
- Turner, S., Hawkesworth, C., Rogers, N. & King, P. U-Th isotope disequilibria and ocean island basalt generation in the Azores. *Chem. Geol.* **139**, 145–164 (1997).
- Schilling, J.-G. *et al.* Petrologic and geochemical variations along the mid-Atlantic ridge from 29°N to 73°N. *Am. J. Sci.* **283**, 510–586 (1983).
- Dosso, L. *et al.* Geochemical structure of the mid-Atlantic ridge between 10° and 41°N. *Earth Planet. Sci. Lett.* **170**, 269–286 (1999).

- Hawkesworth, C. J., Norry, M. J., Roddick, J. C. & Vollmer, R. $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios from the Azores and their significance in LIL-element enriched mantle. *Nature* **280**, 28–31 (1979).
- White, W. M., Tapia, M. D. M. & Schilling, J.-G. The petrology and geochemistry of the Azores Islands. *Contrib. Mineral. Petrol.* **69**, 201–213 (1979).
- Widom, E. & Shirey, S. B. Os isotope systematics in the Azores: implications for mantle plume sources. *Earth Planet. Sci. Lett.* **142**, 451–466 (1996).
- Schiano, P., Birk, J.-L. & Allegre, C. J. Osmium-strontium-neodymium-lead isotopic covariation in mid-ocean ridge basalt glasses and the heterogeneity of the upper mantle. *Earth Planet. Sci. Lett.* **150**, 363–379 (1997).
- Shirey, S. B. & Walker, R. J. The Re-Os isotope system in cosmochemistry and high-temperature geochemistry. *Annu. Rev. Earth Planet. Sci.* **26**, 423–500 (1998).
- Pearson, D. G., Carlson, R. W., Shirey, S. B., Boyd, F. R. & Nixon, P. H. Stabilisation of Archaean lithospheric mantle: A Re-Os isotope study of peridotite xenoliths from the Kaapvaal craton. *Earth Planet. Sci. Lett.* **134**, 341–357 (1995).
- Schaefer, B. F. *et al.* Re-Os isotope characteristics of post-orogenic volcanics: implications for the nature of non-cratonic lithospheric mantle and its contribution to basaltic magmas. *Geology* **28**, 563–566 (2000).
- Bonatti, E. Not so hot “hot spots” in the oceanic mantle. *Science* **250**, 107–111 (1990).
- Jordon, T. H. Composition and development of the continental tectosphere. *Nature* **257**, 745–750 (1978).
- Boyd, F. R. Compositional distinction between oceanic and cratonic lithosphere. *Earth Planet. Sci. Lett.* **96**, 15–26 (1989).
- Campbell, I. H. Implications of Nb/U, Th/U and Sm/Nd in plume magmas for the relationship between continental and oceanic crust formation and the development of the depleted mantle. *Geochim. Cosmochim. Acta* **66**, 1651–1661 (2002).
- Shirey, S. B. & Walker, R. J. Carius tube digestion for low blank rhenium-osmium analysis. *Anal. Chem.* **67**, 2136–2141 (1995).
- Cohen, A. S. & Waters, F. G. Separation of osmium from geological materials by solvent extraction for analysis by thermal ionisation mass spectrometry. *Anal. Chim. Acta* **332**, 269–275 (1996).
- Bennett, V. C., Esat, T. M. & Norman, M. D. Two mantle plume components in Hawaiian picritic lavas from Os-Pb isotopes. *Nature* **381**, 221–224 (1996).
- Hauri, E. H. & Hart, S. R. Re-Os isotope systematics in HIMU and EMII ocean island basalts. *Earth Planet. Sci. Lett.* **114**, 353–371 (1993).
- Hauri, E. H., Lassiter, J. C. & DePaolo, D. J. Osmium isotope systematics of drilled lavas from Mauna Loa, Hawaii. *J. Geophys. Res.* **B 101**, 11793–11806 (1996).
- Hauri, E. H. & Kurz, M. D. Melt migration and mantle chromatography, 2: a time series Os isotope study of Mauna Loa volcano. *Hawaii. Earth Planet. Sci. Lett.* **153**, 21–36 (1997).
- Lassiter, J. C. & Hauri, E. H. Osmium-isotope variations in Hawaiian lavas: evidence for recycled oceanic lithosphere in the Hawaiian plume. *Earth Planet. Sci. Lett.* **164**, 483–496 (1998).
- Marcantonio, F., Zindler, A., Elliott, T. & Staudigel, H. Os isotope systematics of La Palma, Canary Islands — Evidence for recycled crust in the mantle source of himu ocean islands. *Earth Planet. Sci. Lett.* **133**, 397–410 (1995).
- Reisberg, L. *et al.* Os isotope systematics in ocean island basalts. *Earth Planet. Sci. Lett.* **120**, 149–167 (1993).
- Schiano, P. *et al.* Correlated Os-Pb-Nd-Sr isotopes in the Austral-Cook chain basalts: the nature of mantle components in plume sources. *Earth Planet. Sci. Lett.* **186**, 527–537 (2001).

Acknowledgements We thank W. White for supplying samples from Flores and Corvo, M. Gilmour for assistance with the analytical work, and A. Hofmann for comments on the manuscript. This work was supported by the NERC (N.W.R. C.J.H. and I.J.P.) and by a Royal Society university research fellowship (S.P.T.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.T. (e-mail: simon.turner@bristol.ac.uk).

The morphogenesis of feathers

Mingke Yu, Ping Wu, Randall B. Widellitz & Cheng-Ming Chuong

Department of Pathology, Keck School of Medicine, University of Southern California, 2011 Zonal Avenue, Los Angeles, California 90033, USA

Feathers are highly ordered, hierarchical branched structures^{1,2} that confer birds with the ability of flight^{3–5}. Discoveries of fossilized dinosaurs in China bearing ‘feather-like’ structures have prompted interest in the origin and evolution of feathers^{6–14}. However, there is uncertainty about whether the irregularly branched integumentary fibres on dinosaurs such as *Sinornithosaurus* are truly feathers¹¹, and whether an integumentary appendage with a major central shaft and notched edges is a non-avian feather or a proto-feather^{8–10}. Here, we use a developmental approach to analyse molecular mechanisms in feather-branching morphogenesis. We have used the replication-competent avian sarcoma retrovirus¹⁵ to deliver exogenous genes to regenerating flight feather follicles of chickens. We show that the antagonistic balance between noggin and bone morphogenetic protein 4 (BMP4) has a critical role in feather branching, with BMP4 promoting rachis formation and barb fusion, and noggin enhancing rachis and barb branching. Furthermore, we show that sonic hedgehog (Shh) is essential for inducing apoptosis of the marginal plate epithelia, which results in spaces between barbs. Our analyses identify the molecular pathways underlying the topological transformation of feathers from cylindrical epithelia to the hierarchical branched structures, and provide insights on the possible developmental mechanisms in the evolution of feather forms.

With three levels of branching (that is, from rachis to barbs; from

barbs to barbules; and from barbules to cilia or hooklets¹; Fig. 1a) feathers can develop into a variety of forms, including downy feathers, contour feathers, or flight feathers (Fig. 1b). As in hairs, the feather follicle is composed of a dermal papilla and epidermal collar (equivalent to the hair matrix, Fig. 1c–f). Through epithelial–mesenchymal interactions, the epithelial cells at the bottom of the follicle undergo active proliferation (proliferation zone, Fig. 1c). Immediately above this zone, the epithelial cells start to form the rachidial ridge and the barb ridges (ramogenic zone, Fig. 1c)^{16–19}. In a more distal position along the follicle, the barb ridge epithelia actively proliferate and differentiate to form the marginal plates, barbule plates and axial plates (Fig. 1g). The barb ridges grow to form barbs, composed of the ramus and barbules, whereas the marginal and axial plate cells die to become the intervening space. Individual barbule plate cells undergo further cell shape changes to form the cilia and hooklets¹. The barb ridges fuse proximally to form the rachidial ridge, which eventually becomes the rachis. Additional cross-sections (Supplementary Fig. 1) illustrate this process.

The cellular and molecular mechanisms of epithelial organ morphogenesis are beginning to be understood^{20,21}. Although branching morphogenesis²¹ has been studied in the lung and kidney, branching in the feather is unique owing to its exquisite order and non-randomness. Here, we studied the role of noggin–BMP interactions that underlie the fundamental morphogenetic mechanisms^{22–24} in this process. We first analysed the dynamic expressions of BMP2, BMP4 and noggin in remiges (flight feathers) of 15-day-old chicken embryos (E15) using *in situ* hybridization. BMP4 transcripts were detected in the dermal papilla and overlying pulp area (Fig. 1d). At a later time point, BMP4 was expressed in barbule plate cells (Supplementary Fig. 2). BMP2 was present in the marginal plate epithelia in early ramogenesis²⁵, but quickly switched to barbule plate epithelia (Fig. 1f, h, i). BMP4 expression in the

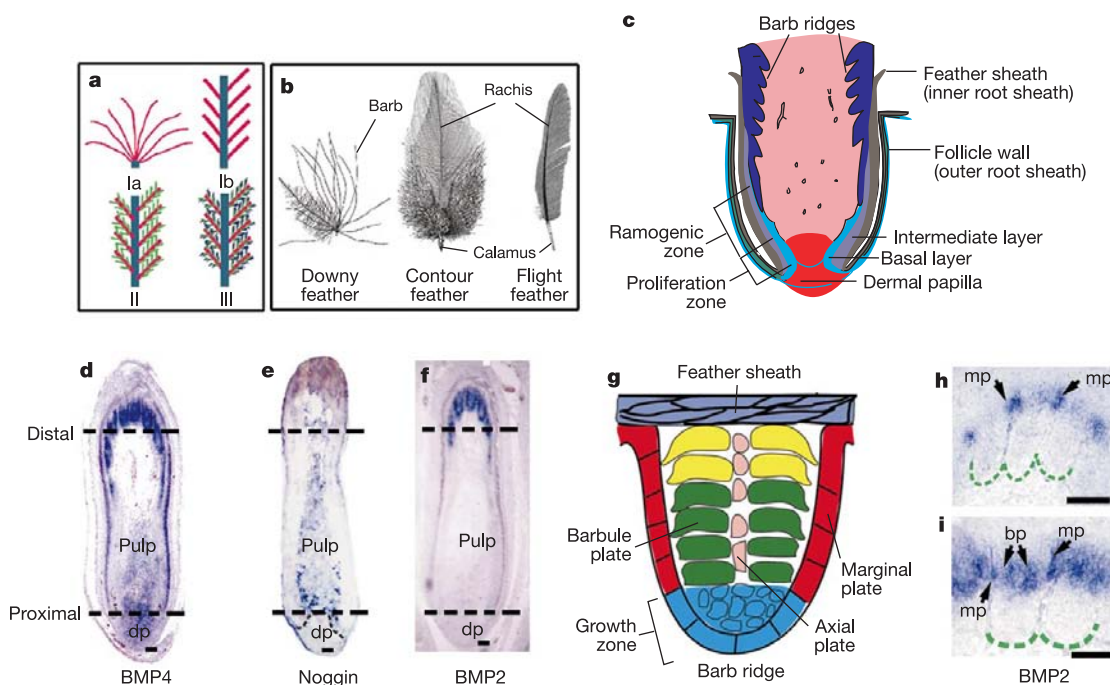


Figure 1 Feather-branching morphogenesis and gene expression. **a**, Diagram showing three branching levels. Level I, rachis (blue) branches into barbs (red). Ia, radially and Ib, bilaterally symmetric feathers. Level II, barbs branch into barbules (green); level III, barbules branch into cilia and hooklets (purple). **b**, Different types of chicken feather. **c**, Diagram of feather follicle structure. **d–f**, BMP4 (**d**), noggin (**e**) and BMP2 (**f**) expression

patterns. The two dotted lines indicate the level of cross-sections shown in Supplementary Fig. 2. **g**, Diagram of feather barb ridge. **h, i**, BMP2 in barb ridges. BMP2 is expressed first in peripheral marginal plates (mp; **h**) then switches to barbule plates (bp; **i**). dp, dermal papilla. Scale bar, 100 μ m.

mesenchyme appeared to form a gradient, tapering from the proximal to distal regions (Fig. 1d). Noggin transcripts were detected in the pulp cells, overlapping with BMP4 transcripts. Noggin was not expressed in the dermal papilla, but was expressed in the pulp regions adjacent to the epidermis. The expression level of noggin appeared to form a gradient from the proximal to distal pulp, with highest expression at the level of the ramogenic zone (Fig. 1e). Cross-sections at the indicated locations (dotted lines) are also shown (Supplementary Fig. 2).

A distinct feature of feathers is that they can regenerate repetitively after plucking. Remige feathers regenerate at a rate of about 0.5 cm day^{-1} , making them excellent recipients for replication-competent avian sarcoma retrovirus (RCAS)-mediated gene expression, which only transduces cells undergoing active mitosis¹⁵. By injecting RCAS retroviral constructs into chick flight feather follicles after plucking (Fig. 2a), exogenous genes were misexpressed during feather regeneration (Fig. 2b, Supplementary Fig. 3). Regenerated feathers injected with RCAS–LacZ virus showed no changes in the formation of rachis or barbs (Fig. 2c, g).

RCAS–noggin was injected into the feather follicles to perturb their BMP activity (Fig. 2d, h; $n = 36$). Many of the regenerated feathers were severely stunted, forming few barbs (33%), and even more had the rachis split into two or four mini-rachides (44%). These smaller rachides gradually converged at the proximal end (Fig. 2d). Barbs of the regenerated feathers expressing exogenous noggin were inhibited (50%) and some barbs (11%) branched into two (Fig. 2h). Histological examination showed that the barb ridges formed irregular tree-like structures with extravagant ridge formations (Fig. 3b, f, j). The rachidial ridge was fragmented into smaller ridges (Fig. 3n).

Regenerated feathers expressing ectopic BMPs had phenotypes that were generally opposite to that of feathers expressing exogenous noggin. Many of the regenerated feathers transduced with

RCAS–BMP4 were short in length, did not form barbs or a rachis (50%), and assumed a morphology similar to the calamus (feather shaft) ($n = 16$). Several of the rachides on these regenerated feathers were oversized (25%). Ectopic rachis-like structures were observed (Fig. 2e; 31%) and the barbs were often fused (Fig. 2i; 38%). Histological sections showed that the barb ridges failed to separate (Fig. 3c, g), the barbs fused (Fig. 3k) and the rachidial ridges were oversized (Fig. 3o). The presumptive marginal plate cells became plump (Fig. 3g) and did not die, thus preventing the creation of spaces between barbs (Fig. 3k).

Overexpression of BMP2 had similar phenotypic effects to BMP4. Feathers showed normal growth at first, but then died abruptly and fell off prematurely at about 3 weeks. The feather size was usually smaller and some showed no barb formation. The rachides of the regenerated feathers were enlarged (Fig. 2f; 23%, $n = 26$) and some feathers had ectopic rachides (23%). Some barbs fused with each other or with the rachis at various points, forming bundles (Fig. 2f, j; 77%). Histological sections showed that some barb ridges fused in pairs (Fig. 3h), suggesting that BMP2 may also function in specifying marginal plate fate, given that BMP2 is expressed transiently in peripheral marginal plates (Fig. 1h). Ectopic rachidial ridge-like structures caused by the fusion of barb ridges were observed (Fig. 3l). The rachidial ridge was extremely large, spanning nearly half of the follicle circumference (Fig. 3p). Thus, the phenotypes of regenerated feathers appeared to depend on the expression levels of the transduced genes.

Whereas abnormal barb ridges shared a forked appearance, the identity of barb branching or barb fusion was confirmed by counting the number and spacing of barbs. Control samples had about six barb ridges in the space shown in Fig. 3e–h. Specimens treated with noggin had elaborately branched ridges alternating with mini-ridge forms, but the total number of ridges was unchanged. Specimens treated with BMP2 had only three forked barb ridges, suggesting that fusion occurred among the original six ridges.

Marginal plate cells express *Shh*²⁶ whereas barbule plate cells express feather keratin (Figs 3q, u and 4a). Characterization of the transduced feathers showed that noggin increased branching by increasing the number of *Shh*-positive marginal plate cells (Fig. 3r) and that it perturbed the shape and arrangement of the barbule plate cells that express feather keratin (Fig. 3v). On the other hand, epithelial overexpression of BMP2 and BMP4 altered the fate of marginal plate cells. The cells, which were plump rather than flattened (Fig. 3g, k) did not die, and branches failed to form. In addition, *Shh* was not expressed (Fig. 3s, t). Some feather keratinization still took place in the barbule plate (Fig. 3w, x).

To test further the role of *Shh* in feather branching, we suppressed *Shh* using cyclopamine²⁷ or RCAS–antisense *Shh* in the plucked and regenerating feather model. The two independent reagents gave similar results. The regenerated feathers showed regions where barbs were fused by means of a web-like membrane, thus forming continuous feather vanes (Fig. 4b, c). Cross-sections showed regions with barb ridges that failed to separate because the marginal plate cells failed to disappear (Fig. 4d, e). Suppressing *Shh* produced a similar phenotype as the overexpression of BMP4 (compare Figs 3g and 4e). In control feathers, TdT-mediated dUTP nick end labelling (TUNEL) staining was positive in the marginal epithelial cells, pulp epithelium and feather sheath (Fig. 4f). The death of these cells allows feather branches to open. Suppression of *Shh* activity increased the number of marginal plate cells. These cells were plump in appearance and were mostly TUNEL negative (Fig. 4g). Thus, overexpression of BMP suppressed *Shh* expression and the subsequent formation of the marginal plate. Suppression of BMP promoted branching, probably by enhancing ridge-forming activity of the basilar cells, together with specifying the fate of marginal plate cells. Therefore *Shh* is required for specification of the fate of marginal plate cells. Balancing the antagonistic actions of *Shh* and

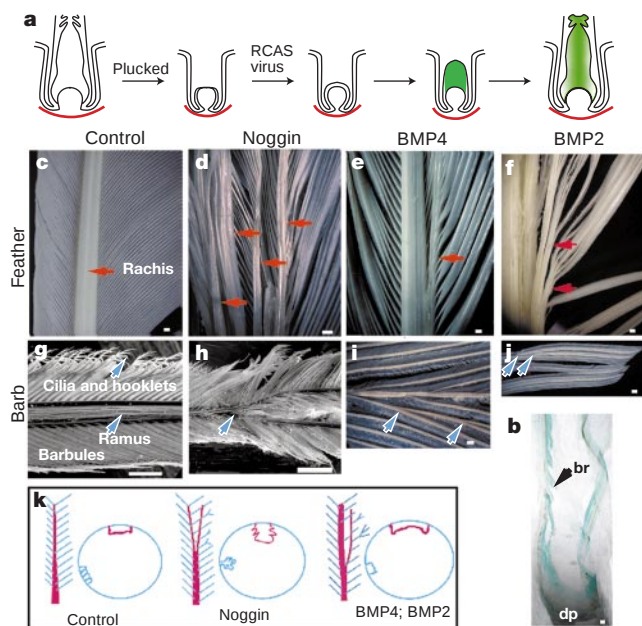


Figure 2 Phenotypic changes in feathers regenerated from follicles injected with RCAS–noggin, RCAS–BMP4 and RCAS–BMP2, respectively. **a**, Diagram showing gene expression strategy. **b**, X-gal staining of a regenerating feather follicle 7 days after RCAS–LacZ injection. **c–f**, Splitting or merging of the rachis is indicated by red arrows. **g–j**, Alteration of the barbs. Abnormal branch points are indicated by blue arrows. **k**, Diagram illustrating the overall phenotypic changes. br, barb ridge; dp, dermal papilla. Scale bar: **b–f, i, j**, 200 μm ; **g, h**, 100 μm .

BMPs sets the number and spacing of barb ridges; however, regulation of Shh by other molecules is also possible.

The importance of the Shh–BMP ‘signalling module’²⁸ in skin appendage morphogenesis has also been shown by comparing chicken feather variants, duck feathers, and avian and alligator scales²⁵. Using embryonic chicken explant cultures, this study indicated that Shh is important for proliferation in proximal barb ridges, whereas BMP2 suppresses Shh and promotes differentiation of distal barb ridges²⁵. Our *in vivo* ‘transgenic feather’ model using plucked/regenerated mature feather follicles allows us to analyse the late-branching event that is impossible to observe in embryonic explants. This model should make it possible for future experiments to link molecular pathways with feather forms.

On the basis of these data, we propose the following model for feather-branching morphogenesis (Fig. 5a, b). The multilayered epidermis can be moulded into epidermal ridges by ridge-forming activators (for example, noggin) and inhibitors (for example, BMPs) distributed in the adjacent mesenchyme. The balance between noggin and BMP4 determines the number, size and spacing of barb ridges. First, at the proliferation zone, the level of BMP4 is much higher than that of noggin, thus the epithelial cells form a cylindrical structure. Second, at the ramogenic zone, the level of

noggin in the pulp adjacent to epithelia gradually exceeds that of BMP4, thus the epithelial cells start to form multiple barb ridges. Third, the basilar layer becomes periodically arranged into Shh positive/BMP2 transiently positive marginal plate cells that die and Shh negative cells from the barb ridge growth zone that proliferate. The intermediate cell layer is ‘cleaved’ into groups of cells that become the initial barb ridge. Marginal plate cells die to ensure the formation of spaces between barb ridges^{1,29}. Fourth, the originally randomly arranged cells in the barb ridge express BMP2 and BMP4, line up, and become two rows of barbule cells. These cells further differentiate to form the barbules. Barbule plate cells¹ stimulated by BMP2 and BMP4 change shape to form the cilia and hooklets. Fifth, towards the end of feather formation, noggin activity is reduced and conditions revert back to stage one, forming the calamus without branches at the proximal end of the feather shaft. Last, if the noggin:BMP ratio becomes polarized in the anterior–posterior axis, the site with higher BMP activity eventually becomes the rachis, thus the bilaterally symmetric feather can form.

We have demonstrated that the BMP pathway regulates the formation and inter-conversion of the barbs and rachis. We have also shown that the true separation of branches requires Shh activity. Other morphogens, such as fibroblast growth factors

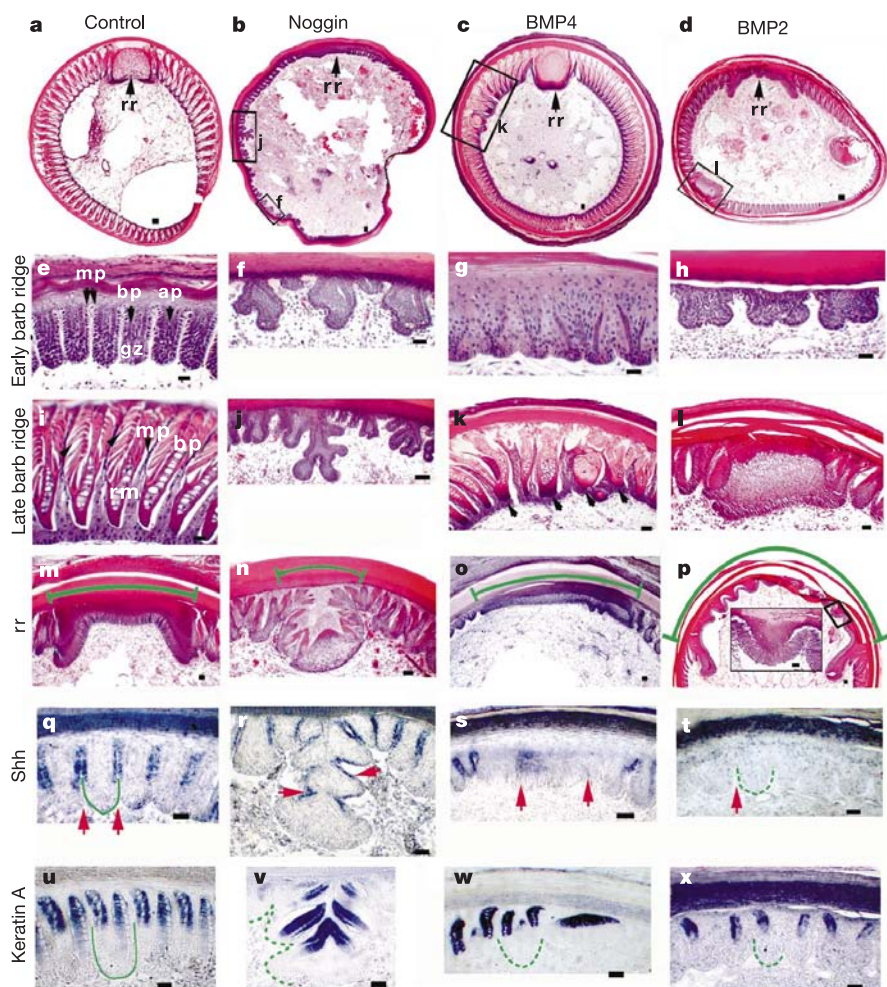


Figure 3 Analyses of feathers injected with RCAS–noggin, RCAS–BMP4 and RCAS–BMP2, respectively. **a–d**, Cross-sections of a feather. **e–l**, Changes in barb ridges. **m–p**, Changes of rachis width (stained with haematoxylin and eosin). The rachis width is indicated by a green arc. The inset in **p** is part of the rachis structure, not barb ridges (compare with **e**). **q–t**, *In situ* hybridization with Shh probes. Marginal plates or

comparable regions are indicated with red arrows. **u–x**, *In situ* hybridization with feather keratin probes. Green lines delineate barb ridges; dotted lines indicate abnormal barb ridges. ap, axial plate; bp, barbule plate; gz, barb ridge growth zone; mp, marginal plate; rr, rachis ridge. Scale bar, 100 μ m.

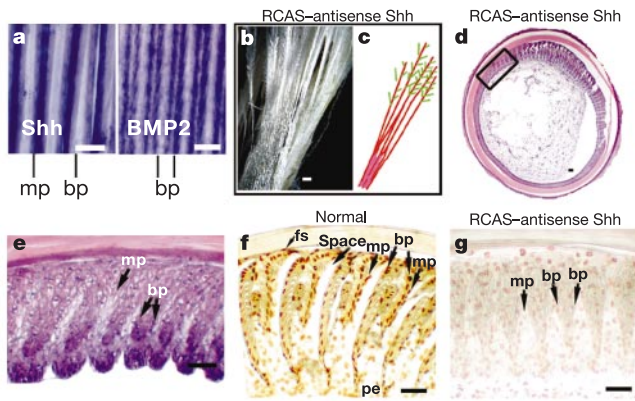


Figure 4 The role of Shh in barb formation. **a**, Normal feather whole-mount *in situ* hybridization of Shh and BMP2. **b**, Shh inhibition by RCAS-antisense Shh or cyclopamine produces fused vanes. **c**, Diagram showing changes in **b**. Barbs are red and barbules are green. **d**, Cross-sections (stained with haematoxylin and eosin) show barb ridge segments that fail to separate, varying from more severe (bottom left quarter of cross-section) to less severe (top right segment). **e**, Greater magnification of boxed section in **d**. **f, g**, Apoptosis in late-differentiated barb ridges in normal and Shh-suppressed follicles. bp, barbule plate; fs, feather sheath; mp, marginal plate; pe, pulp epithelium. Scale bar, 100 μ m.

(FGFs) and Wnts¹⁸, may also be involved in this process, and we expect them to behave under similar principles along the same pathway or with special regulation to produce feather variants.

Formation of hierarchical branches is the principal feature of feathers¹⁷, and is therefore one of the chief issues in the origin and evolution of feathers. On the basis of some fossil evidence it has been proposed that a filamentous integument structure with a major central shaft and notched edges may be the prototype of feathers^{8–10}. According to this model, the rachis would have formed first in evolution, then barbs, and finally barbules. Therefore, the rachis and barbs would be different entities and not interchangeable (Fig. 5c). Alternatively, because barbs form first during development, it was proposed that barbs appeared first in integument evolution, and the rachis, a specialized form of fused barbs, appeared later as an evolutionary novelty^{16,18}. The fact that the barbs and the rachis

can be converted experimentally in the laboratory favours the barb to rachis model. Our data suggest that a radially symmetric feather is more primitive than the bilaterally symmetric feather in terms of molecular and developmental mechanisms, and may have been the prototype of feathers (Fig. 5c). Some fossilized primitive skin appendages on *Sinornithosaurus* also favour this model¹¹. Further modulation of BMP and Shh pathways may have led to the many varieties of feather seen today by regulating the number, shape and size of the rachis, barbs, and barbules^{1,17,30}. This work provides evidence for the molecular mechanisms possibly involved in the evolution of feather branching. □

Methods

Materials

Specific pathogen-free, fertilized white leghorn chicken eggs were purchased from Charles River SPAFAS. The eggs were incubated at 38 °C in a humidified rotating incubator. Chicks were housed in the University of Southern California vivarium. Digoxigenin-labelled probes were generated by *in vitro* transcription from plasmids provided by L. Niswander for noggin, P. Francis-West for BMP2 and BMP4, and C. Tabin for Shh. A probe for feather keratin was prepared in our laboratory, GenBank accession number X 17511. RCAS-noggin plasmid was from R. Johnson; RCAS-BMP2 and RCAS-BMP4 plasmids were from P. Francis-West; RCAS-LacZ plasmid encoding β -galactosidase was originally constructed by L. Yi and was provided by W.-P. Wang; RCAS-Shh sense was from C. Tabin; RCAS-Shh antisense was produced by cutting the sense Shh plasmid with *Cla*I and re-ligating the inserted Shh in the reverse orientation. The orientation was confirmed by restriction enzyme analysis (S. A. Ting-Bereth and C.-M.C., unpublished data). Viruses were prepared and titrated²⁴. Cyclopamine, a Shh antagonist²⁶ provided by W. Gaffield, was dissolved in absolute ethanol and diluted to 0.4 mg ml⁻¹ in DMEM for injection (20 μ l) into the regenerating feathers, using a similar procedure described for viral transduction.

Transduction of regenerating feather follicles

Chickens at 2–3 weeks of age were anaesthetized with ketamine (50 mg per kg body mass). Primary remiges I to VII were used because of their large size, distinct morphology and identity. Normally, feathers will regenerate and grow out of the follicles at about 14 days after plucking. Regenerated feathers/feather follicles were dissected at 7 days after plucking. Serial longitudinal and cross paraffin sections were cut and stained with haematoxylin and eosin. For gene transduction, RCAS viruses were injected into the empty follicles immediately after feathers were plucked. RCAS virus propagation in follicles was verified using RCAS-LacZ virus followed by 5-bromo-4-chloro-3-indolyl- β -D-galactoside (X-gal) staining of cryostat sections at various days after the injection. Each follicle received about 10–20 μ l of medium containing the virus (1×10^5 to 1×10^6 infectious units per ml). For preliminary studies, viruses containing the genes of interest were injected into primary remige follicles on the left wing. RCAS-LacZ virus was injected into primary remige follicles on the right wing as controls. Feather follicles injected with viruses were dissected and sectioned for histological study at 2–3 weeks after virus injection. Once distinct phenotypes were confirmed in at least three repeat experiments,

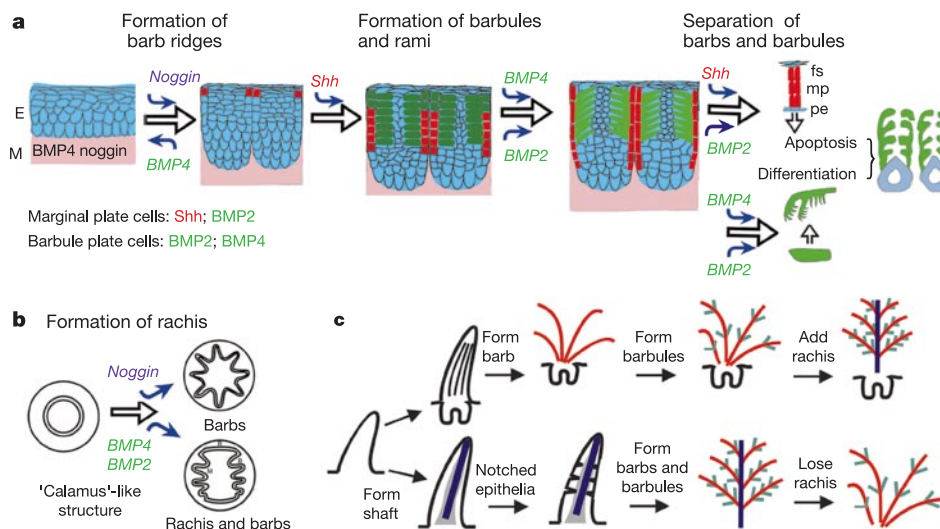


Figure 5 Models of feather branching and evolution of feather forms. **a**, Roles of noggin/BMP4, Shh and BMP2 in the three levels of feather branching. E, epithelial cells (blue); M, mesenchymal cells (pink). fs, feather sheath; mp, marginal plate; pe, pulp epithelium. **b**, The ratio of noggin and BMP4 may determine the number and size of barb ridges. A

localized high BMP:noggin ratio, together with a helical growth mode of barb ridges¹⁷, can lead to the formation of a rachidial ridge through fusion of barb ridges. **c**, Hypothetical models of the evolution of feather forms. Top row, barb to rachis model; bottom row, rachis to barb model. The experimental data are in favour of the barb to rachis model.

viruses were injected to follicles on both wings for later studies. Chickens were raised in cages and observed on a daily basis over a two-month period. The regenerated feathers were plucked and examined with a dissection or scanning electron micrograph microscope for abnormalities compared with normal primary remiges.

Histology and *in situ* hybridization

Paraffin sections (5 µm) were stained with haematoxylin and eosin or prepared for *in situ* hybridization following routine procedures²⁶. Cryostat sections (10 µm) were stained with X-gal. TUNEL staining was performed using a kit (Roche). Nonradioactive wholemount or section *in situ* hybridization or section *in situ* hybridization was performed according to the protocol described^{22,26}. After hybridization, sections were incubated with an anti-digoxigenin Fab conjugated to alkaline phosphatase (Boehringer Mannheim). Colour was detected by incubating with a Boehringer Mannheim purple substrate (Roche).

Received 5 June; accepted 10 October 2002; doi:10.1038/nature01196.

Published online 30 October 2002.

- Lucas, A. M. & Stettenheim, P. R. (eds) *Avian Anatomy – Integument*. Agricultural Handbook 362: Agricultural Research Services (US Department of Agriculture, Washington DC, 1972).
- Chuong, C.-M. The making of a feather: Homeoproteins, retinoids and adhesion molecules. *BioEssays* **15**, 513–521 (1993).
- Feduccia, A. *The Origin and Evolution of Birds* 2nd edn (Yale Univ. Press, New Haven, Connecticut, 1999).
- Chatterjee, S. *The Rise of Birds* (John Hopkins Univ. Press, Baltimore, Maryland, 1997).
- Regal, P. J. The evolutionary origin of feathers. *Q. Rev. Biol.* **50**, 35–66 (1975).
- Chen, P. J., Dong, Z. M. & Shen, S. N. An exceptionally well-preserved theropod dinosaur from the Yixian Formation of China. *Nature* **391**, 147–152 (1998).
- Xu, X., Tang, Z. L. & Wang, X. L. A therizinosaurid dinosaur with integumentary structures from China. *Nature* **399**, 350–354 (1999).
- Jones, T. D. *et al.* Nonavian feathers in a late Triassic archosaur. *Science* **288**, 2202–2205 (2000).
- Prum, R. O. Longisquama fossil and feather morphology. *Science* **291**, 1899–1902 (2001).
- Zhang, F. & Zhou, Z. A primitive enantiornithine bird and the origin of feathers. *Science* **290**, 1955–1959 (2000).
- Xu, X., Zhou, Z. & Prum, R. O. Branched integumental structures in Sinornithosaurus and the origin of feathers. *Nature* **410**, 200–204 (2001).
- Ji, Q., Currie, P. J., Norell, M. A. & Ji, S. A. Two feathered dinosaurs from northeast China. *Nature* **393**, 753–761 (1998).
- Ji, Q., Norell, M. A., Gao, K. Q., Ji, S. A. & Ren, D. The distribution of integumentary structures in a feathered dinosaur. *Nature* **410**, 1084–1088 (2001).
- Norell, M. *et al.* Modern feathers on a non-avian dinosaur. *Nature* **416**, 36–37 (2002).
- Morgan, B. A. & Fekete, D. M. Manipulating gene expression with replication-competent retroviruses. *Methods Cell Biol.* **51**, 185–218 (1996).
- Prum, R. O. Development and evolutionary origin of feathers. *J. Exp. Zool.* **285**, 291–306 (1999).
- Prum, R. O. & Williamson, S. Theory of the growth and evolution of feather shape. *J. Exp. Zool.* **291**, 30–57 (2001).
- Chuong, C.-M., Chodankar, R., Widelitz, R. B. & Jiang, T.-X. Evo-devo of feathers and scales: building complex epithelial appendages. *Curr. Opin. Genet. Dev.* **10**, 449–456 (2000).
- Chuong, C.-M. *et al.* Dinosaur's feather and Chicken's tooth? Tissue engineering of the integument. John Ebling lecture. *Eur. J. Dermatology* **11**, 286–292 (2001).
- Chuong, C.-M. (ed.) *Molecular Basis of Epithelial Appendage Morphogenesis* (Landes Bioscience, Austin, 1998).
- Hogan, B. L. M. Morphogenesis. *Cell* **96**, 225–233 (1999).
- Jung, H.-S. *et al.* Local inhibitory action of BMPs and their relationships with activators in feather formation: implications for periodic patterning. *Dev. Biol.* **196**, 11–23 (1998).
- Dudley, A. T. & Tabin, C. J. Constructive antagonism in limb development. *Curr. Opin. Genet. Dev.* **10**, 387–392 (2000).
- Jiang, T.-X., Jung, H.-S., Widelitz, R. B. & Chuong, C.-M. Self organization of periodic patterns by dissociated feather mesenchymal cells and the regulation of size, number and spacing of primordia. *Development* **126**, 4997–5009 (1999).
- Harris, M. P., Fallon, J. F. & Prum, R. O. Shh-Bmp2 signaling module and the evolutionary origin and diversification of feathers. *J. Exp. Zool.* **294**, 160–176 (2002).
- Ting-Berthel, S. A. & Chuong, C.-M. Sonic hedgehog in feather morphogenesis: induction of mesenchymal condensation and association with cell death. *Dev. Dyn.* **207**, 157–170 (1996).
- Cooper, M. K., Porter, J. A., Young, K. E. & Beachy, P. A. Teratogen-mediated inhibition of target tissue response to Shh signaling. *Science* **280**, 1603–1607 (1998).
- Calabretta, R., Nolfi, S., Parisi, D. & Wagner, G. P. Duplication of modules facilitates the evolution of functional specialization. *Artificial Life* **6**, 69–84 (2000).
- Chuong, C.-M. & Edelman, G. M. Expression of cell adhesion molecules in embryonic induction. II. Morphogenesis of adult feathers. *J. Cell Biol.* **101**, 1027–1043 (1985).
- Gill, F. B. *Ornithology*, 2nd edn (Freeman, New York, 1994).

Supplementary Information accompanies the paper on Nature's website
(<http://www.nature.com/nature>).

Acknowledgements We thank M. Ramos for help in preparing the manuscript; and R. Prum for critical comments on the manuscript. Figure 1b is modified from ref. 1. This work is supported by grants from the National Institute of Arthritis and Musculoskeletal and Skin Diseases, USA, and the National Science Foundation to C.-M.C., and a National Cancer Institute grant to R.B.W.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to C.-M.C.
(e-mail: chuong@pathfinder.usc.edu).

The genome sequence and structure of rice chromosome 1

Takuji Sasaki*, Takashi Matsumoto*, Kimiko Yamamoto*, Katsumi Sakata*, Tomoya Baba*, Yuichi Katayose*, Jianzhong Wu*, Yoshihito Niimura†, Zhukuan Cheng‡, Yoshiaki Nagamura*, Baltazar A. Antonio*, Hiroyuki Kanamori*, Satomi Hosokawa*, Masatoshi Masukawa*, Koji Arikawa*, Yoshino Chiden*, Mika Hayashi*, Masako Okamoto*, Tsuyu Ando*, Hiroyoshi Aoki*, Kohei Arita*, Masao Hamada*, Chizuko Harada*, Saori Hijishita*, Mikiko Honda*, Yoko Ichikawa*, Atsuko Itonuma*, Masumi Iijima*, Michiko Ikeda*, Maiko Ikeno*, Sachie Ito*, Tomoko Ito*, Yuichi Ito*, Yukiyo Ito*, Aki Iwabuchi*, Kozue Kamiya*, Wataru Karasawa*, Satoshi Katagiri*, Ari Kikuta*, Noriko Kobayashi*, Izumi Kono*, Kayo Machita*, Tomoko Maehara*, Hiroshi Mizuno*, Tatsumi Mizubayashi*, Yoshiyuki Mukai*, Hideki Nagasaki*, Marina Nakashima*, Yuko Nakama*, Yumi Nakamichi*, Mari Nakamura*, Nobukazu Namiki*, Manami Negishi*, Isamu Ohta*, Nozomi Ono*, Shoko Saji*, Kumiko Sakai*, Michie Shibata*, Takanori Shimokawa*, Ayahiko Shomura*, Jianyu Song*, Yuka Takazaki*, Kimihiro Terasawa*, Kumiko Tsuji*, Kazunori Waki*, Harumi Yamagata*, Hiroko Yamane*, Shoji Yoshiki*, Rie Yoshihara*, Kazuko Yukawa*, Huisun Zhong*, Hisakazu Iwama†, Toshinori Endo§, Hidetaka Ito§, Jang Ho Hahn||, Ho-Il Kim||, Moo-Young Eun||, Masahiro Yano*, Jiming Jiang‡ & Takashi Gojobori†

* Rice Genome Research Program, National Institute of Agrobiological Sciences, and Institute of the Society for Techno-innovation of Agriculture, Forestry and Fisheries, 1-2, Kannondai 2-chome, Tsukuba, Ibaraki 305-8602, Japan

† Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, Mishima 411-8540, Japan

‡ Department of Horticulture, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

§ Department of Bioinformatics, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan

|| Rice Genome Sequencing Project, National Institute of Agricultural Science and Technology, RDA, 249 Seodun-dong, Suwon 441-707, Korea

The rice species *Oryza sativa* is considered to be a model plant because of its small genome size, extensive genetic map, relative ease of transformation and synteny with other cereal crops^{1–4}. Here we report the essentially complete sequence of chromosome 1, the longest chromosome in the rice genome. We summarize characteristics of the chromosome structure and the biological insight gained from the sequence. The analysis of 43.3 megabases (Mb) of non-overlapping sequence reveals 6,756 protein coding genes, of which 3,161 show homology to proteins of *Arabidopsis thaliana*, another model plant. About 30% (2,073) of the genes have been functionally categorized. Rice chromosome 1 is (G + C)-rich, especially in its coding regions, and is characterized by several gene families that are dispersed or arranged in tandem repeats. Comparison with a draft sequence⁵ indicates the importance of a high-quality finished sequence.

Rice has been studied extensively by molecular genetics and constitutes one of the best characterized crop plants with a fine genetic map of 3,267 markers (<http://rgp.dna.affrc.go.jp/public-data/geneticmap2000/index.html>)¹, a yeast artificial chromosome (YAC) physical map with 80.8% coverage², sequences for about 10,000 unique expressed sequence tags (ESTs)³, and a transcriptional map indicating the placement of 6,591 unique ESTs². The Rice Genome Research Program (RGP) in Japan launched its rice genome sequencing project in 1998. It is a partner of the International Rice Genome Sequencing Project (IRGSP), which involves ten countries in Asia, North America, South America and Europe that are working towards the immediate release of high-quality sequence data to the public domain⁴. The draft sequences of the two

main subspecies of rice, *japonica* and *indica*, have been reported^{5,6}. Both studies were based on whole-genome shotgun sequencing rather than on the clone-by-clone approach of the IRGSP. Although the release of the draft sequence is of immense scientific value, many challenges in rice genomics demand the availability of a complete, accurate, map-based rice genome sequence.

We determined the sequence of chromosome 1 from 390 overlapping phage (P1)-derived artificial chromosome (PAC) and bacterial artificial chromosome (BAC) clones and assembled it into nine contigs (Fig. 1). The longest contig is 14.4 Mb and spans positions 106.2 centimorgans (cM) to 157.1 cM on the molecular genetic map. Among the eight remaining gaps, gap 4, located at 73.4 cM, corresponds to a portion of the centromeric region and is estimated to be about 1,400 kilobases (kb) by the pachytene fluorescence *in situ* hybridization (FISH) method⁷. PAC/BAC clones adjacent to this gap contain copies of the rice centromere-specific sequence RCS2 (ref. 8). Two PAC clones, P0402A09 and P0020E09, are localized to the most distal ends of the short arm and the long arm, and their map positions have been verified by pachytene FISH using pAtT4 (ref. 9), a telomeric clone of *Arabidopsis* (Supplementary Fig. 1). This indicates that our physical map extends to within less than 50 kb of the telomeres. Integration of the PAC/BAC physical mapping with the results from fibre FISH gives a total length of 45.7 Mb for chromosome 1, corresponding to 181.8 cM on the genetic map, excluding the telomeres.

Statistics for the nucleotide sequence of rice chromosome 1 are summarized in Table 1. The non-overlapping sequence covers 43,276,883 nucleotides. In this sequence, 6,756 genes were either identified or predicted. Thus, the average gene density of chromosome 1 is about one gene per 6.4 kb. If this distribution is assumed to be similar throughout the whole genome, then the total number of genes in the rice genome (400 Mb) is roughly 62,500. This number is 2.5 times larger than the gene total of *Arabidopsis*¹⁰. But this difference might easily be the result of an overestimate of rice genes, because it assumes that there is a uniform distribution of genes along the chromosomes.

Cytogenetic analysis has indicated clear differences in the content of heterochromatin in each of the 12 rice chromosomes, and chromosome 1 shows the least amount of heterochromatic

Table 1 Compositional analysis of the sequence of rice chromosome 1

Overall physical length	45,745,883 bp
Short arm	17,081,313 bp
Long arm	27,264,570 bp
Total length of eight gaps	2,469,000 bp
Non-overlapping sequence	43,276,883 bp
Base composition (% GC)	
Overall	43.8%
Coding region	58.2%
Noncoding region	40.7%
Predicted gene number	6,756
Gene density	6.4 kb per gene
Average gene size	3.4 kb
Exons	
Size of exons	1.1 kb per gene
Number of exons per gene	4.8
Introns	
Size of introns	2.3 kb per gene
Number of introns per gene	3.8
Repetitive sequences	
Number of retrotransposons (class I)	3,235
Number of DNA transposons (class II)	
Autonomous type	6,985
Nonautonomous type (MITEs)	14,106
Total number of repeats	24,326

material¹¹. The average exon size is comparable to that of *Arabidopsis*, but the average intron size is about 3.6 times larger. This means that, although the longer introns engender larger gene sizes in rice, the average transcriptome size is similar in both species. The G + C content of coding and noncoding regions in rice is higher than in *Arabidopsis*—the rice coding regions are especially (G + C)-rich. This characteristic is reflected by the biased usage of G/C at the third position of codons within predicted genes (Supplementary Table 1). Buoyant density experiments have shown that rice genes are localized in (G + C)-rich islands that occupy 24% of the genome¹². When we plotted the average G + C values against chromosomal position in chromosome 1, however, we did not detect any CpG islands, indicating a neutral nucleotide distribution. The ratios of physical to genetic distance on the short and the long arms are 214 kb cM⁻¹ ($r^2 = 0.983$) and 288 kb cM⁻¹ ($r^2 = 0.976$), respectively, suggesting that the rate of recombination differs along the two arms of the chromosome.

We compared our finished sequence (493,729 bp from the distal

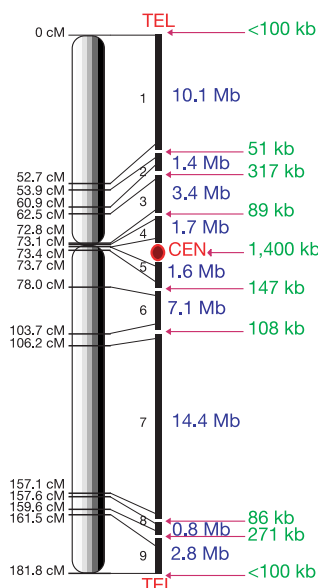


Figure 1 Physical map of rice chromosome 1. Positions of the PAC/BAC contigs are indicated by black bars. Purple numbers indicate the physical distances that were calculated on the basis of the nucleotide sequence length of each contig. A representation of the genetic map of chromosome 1 is shown on the left with the positions of the genetic

markers found nearest the end of each contig. The centromeric region is shown as a red circle. The green numbers show the gap sizes as measured by fibre FISH and pachytene FISH.

end of the short chromosome arm) with 127,550 *indica* sequence contigs assembled from the whole-genome shotgun sequences of the Beijing Genomics Institute (BGI, <http://btn.genomics.org.cn/rice/>) using the *japonica* sequence as a query for basic BLASTN (basic local alignment search tool) analysis (Fig. 2). We could detect the corresponding *indica* sequence in about 78% of the whole region. But there were 65 gaps in the aligned contigs, and a total of 110,389 bases (22%) of *japonica* sequence could not be identified in the *indica* assembly. This may partly reflect the sequence difference between the two subspecies, although some artefacts in the whole-genome shotgun assembly cannot be ruled out. Among the 96 predicted genes in this region of the completed *japonica* sequence, 55 genes are intact, 33 genes are partially predicted and 8 genes are not predicted in the corresponding *indica* draft sequence. Relative identities near the repeat (retrotransposon-like) regions are lower than in the other regions, indicating a misassembly in the sequence.

Direct comparison with the *japonica* draft sequence could not be made because the sequence data are not in the public domain. But previously, 4,467 genes were predicted from a set of 99 BAC contigs assigned to chromosome 1 (ref. 6). It is likely that an estimated 2,835–4,211 gaps (either 63 gaps per megabase or 10% of 42,109 total gaps) for this chromosome prevented an accurate prediction of the number of genes. Not surprisingly, only half of the genes predicted contain complete coding regions. In addition, no basis was provided for the assignment of genes to chromosome⁶.

We used an automated annotation system, RiceGAAS¹³, to characterize the gene composition of chromosome 1 (Supplementary Fig. 2, <http://RiceGAAS.dna.affrc.go.jp/chromosome1/>). The distribution of genes along both arms of the chromosome indicates higher density (18–19 genes per 100 kb) in distal as compared with proximal regions (10–12 genes per 100 kb). This was verified by experimental results obtained by mapping 977 expressed sequence tags on to chromosome 1 (ref. 2). Among the 6,756 predicted genes, 2,073 (31%) were functionally characterized by homology to known proteins using BLASTP, whereas 69% of the predicted genes corresponded to proteins with no known function (Table 2). The protein signature search program InterPro detected protein domains in 3,660 (54%) of the total predicted genes (see <http://RiceGAAS.dna.affrc.go.jp/chromosome1/>). In particular, 1,170 (33%) of 3,600 hypothetical proteins showed domain homology, suggesting that these proteins may correspond to newly identified proteins in rice. BLASTN analysis was done using the cereal EST entries from the EST database at the National Institute for Biotechnology Information (NCBI). Exon regions from all predicted genes

Table 2 Functional classification of the proteins encoded on rice chromosome 1

Total proteins*	6,756
Class I	40
Class II	799
Class III	1,234
Class IV	870
Class V	213
Class VI	3,600
Functional classification	
Cellular metabolism	421 (6.23%)
Signal transduction	365 (5.40%)
Transcription	255 (3.77%)
Cell rescue, defence	215 (3.18%)
Transport facilitation	172 (2.55%)
Cellular organization	132 (1.95%)
Protein destination	108 (1.60%)
Protein synthesis	85 (1.26%)
Energy	83 (1.23%)
Cell growth	62 (0.92%)
Development	35 (0.52%)
Cellular transport	28 (0.41%)
Cellular biogenesis	16 (0.24%)
Classification not yet clear-cut	13 (0.19%)
Ionic homeostasis	9 (0.13%)
Organism-specific	3 (0.04%)
Unclassified	4,754 (70.37%)

* Class I, complete match to a rice protein; class II, strong similarity to a known protein, from the N to the C terminus; class III, similarity to less than 50% of the full length of the target protein; class IV, similarity to a protein of unknown function predicted by genome sequence annotation; class V, match to an EST nucleotide sequence but not to a known protein; class VI, no hit at a threshold probability score of 10^{-20} to any entry in the database.

were used as queries, and 546,723 unclustered ESTs from wheat, maize, barley and sorghum were searched using a threshold probability value of 10^{-5} . A total of 2,985 predicted genes, including 756 hypothetical proteins, have cereal homologues. Thus, among the 6,756 predicted genes, 4,803 (71%) show some evidence of homology to a domain, a functional site, a cereal EST or a protein.

The predicted proteins found on chromosome 1 were categorized into gene families by BLASTP, using a threshold probability score of 10^{-20} over more than 50% of the length of the gene. The most abundant gene family was the serine/threonine receptor kinase family with 132 members distributed along the chromosome (Fig. 3a). A cluster of this gene family was observed at the distal end of the short arm, although some members of the cluster seemed to be pseudogenes. The highest number of tandem repeats detected at a single site was a cluster of ten copies of the hypothetical gene family located on the short arm of chromosome 1. These results are summarized in Fig. 3b, which shows a dot matrix plot of chromosome 1, indicating the predicted genes with significant homology to a given gene. On this plot, which disregards self-homology, a clear diagonal line was obtained, indicating that a significant number of

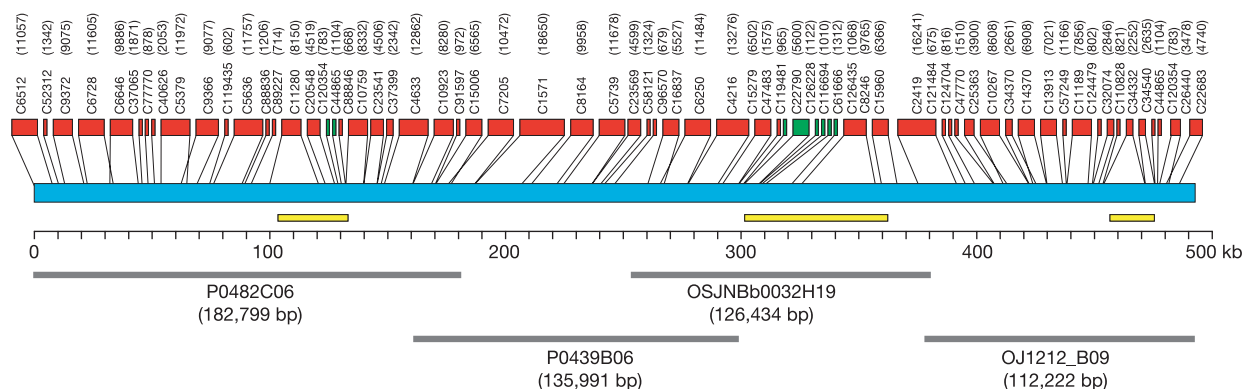


Figure 2 Comparison between the Nipponbare finished sequence and the *indica* draft sequence. Our finished continuous sequence (493,729 bp from near the distal end of the short arm) was used as query. All of the *indica* 93–11 sequence contigs assembled from the whole-genome shotgun sequences from the BGI were downloaded from their website and searched by BLASTN. The highest rank of the hit contigs was aligned to our

Nipponbare sequence. Coloured bars represent the following: blue, RGP sequence data; grey, PAC/BAC clones; red, BGI sequence data with >95% identity; green, BGI sequence data with 90–95% identity; and yellow, regions containing repetitive sequences. Numbers in parentheses correspond to the length of PAC/BAC clones and to BGI sequence contigs in bp.

genes are duplicated and arrayed in tandem.

To determine whether any of the proteins on rice chromosome 1 are not present in *Arabidopsis*, the 6,756 predicted proteins were queried in BLASTP searches against all the *Arabidopsis* proteins in the Munich Information Center for Protein Sequence (MIPS) database using a threshold probability score of 10^{-5} . Among 3,161 positive queries, 824 showed strong similarities (probability value less than 10^{-100}) to proteins found in *Arabidopsis*, whereas 3,595 sequences (53%) did not have positive BLASTP hits with predicted *Arabidopsis* proteins at a probability threshold of 10^{-5} . Only 27 of these sequences had homology to known proteins and among them, only Bowman–Birk trypsin inhibitor and cytochrome *f* (chloroplast) were clearly found in rice chromosome 1. This suggests that almost all of the known proteins found in rice chromosome 1 are also found in *Arabidopsis*. Among the hypothetical proteins, 3,051 genes have no counterpart in *Arabidopsis* and 442 (15%) genes have grass orthologues. Analysis of the draft sequence also showed that half of the predicted genes have no homologues in *Arabidopsis*^{5,6}. Although many of these hypothetical genes could be artefacts resulting from prediction errors, functional characterization of these genes in the future may identify grass-specific or even rice-specific genes.

We also observed rice chloroplast genes in sequential order on the

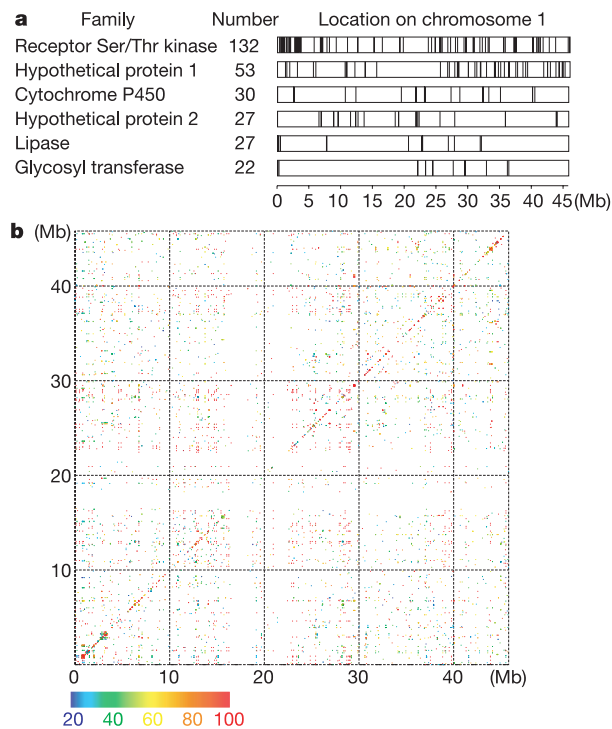


Figure 3 Analysis of gene families and gene clusters. **a**, Distribution of gene families on chromosome 1. For each category, BLASTP was carried out against all of the chromosome 1 gene products. Proteins that had a threshold probability (*E*) value of less than 10^{-20} and that also showed strong homology for more than half of their total length were grouped as a gene family. The six largest gene families are shown. Vertical lines on each bar indicate the positions of gene family members. The bars are oriented with the distal end of the short arm of the chromosome to the left. The scale at the bottom shows the physical length of chromosome 1 in Mb. **b**, Dot matrix plot showing the positions of homologous genes on chromosome 1. A BLASTP search was carried out using all of the predicted genes. A dot was plotted when the *E* value between two genes was less than 10^{-20} and the length of the match was over 50%. The colour of each dot represents the *E* value of the match. The colour spectrum bar at the bottom left shows the *E* value associated with each colour; for example, green corresponds to $E = 10^{-40}$. The matches of $E < 10^{-100}$ are shown as red dots. 'Self-against-self' matches are omitted.

chromosomal DNA. For example, at 149.1 cM we identified 3,564 bp of sequence that matched the rice chloroplast sequence with only a 3-bp difference. This sequence contains three genes¹⁴, PSII cytochrome *b*₅₅₉, cytochrome *f* and the chloroplast envelope membrane protein ORF230. We also detected 85 putative transfer RNA genes using tRNAscan SE¹⁵. Analysis of the retrotransposable elements and DNA intermediate transposons, including miniature inverted-repeats transposable elements (MITEs)¹⁶, using Repeat-Masker is given in Table 1 and summarized in Supplementary Fig. 3. MITEs have a tendency to be dispersed along the chromosome, whereas the retrotransposons and other autonomous type DNA-mediated transposable elements are clustered in the pericentromeric region. Among retroelements, Ty3/*Gypsy*-type elements are the most frequent (2,157), followed by Ty1/*Copia*-type elements (384). The sum of the lengths of these three repetitive elements is 6.0 Mb, corresponding to 13% of chromosome 1.

There are at least three compelling reasons for obtaining finished high-quality sequence for the complete rice genome: first, the ability to determine gene function is highly dependent on having accurate sequences; second, as a model plant for the cereal grasses, the complete rice sequence will directly affect what can be accomplished with the other cereal grasses; and last, the identification of genes responsible for agronomic traits of economic importance requires precise map-based genomic sequence. Chromosome 1 contains many biologically important genes. More than 20 gene loci have been identified by genetic analysis, including genes controlling dwarfing and fertility. One of these genes, *sd1* has been cloned and shown to encode one of the enzymes in gibberellic acid synthesis¹⁷.

The complete genomic sequence of chromosome 1 has yielded several findings that would be observed only using a clone-by-clone sequencing strategy. Gene families comprising active and inactive members and sets of tandemly repeated genes seem to be common features of chromosome 1. This redundancy may account for the unexpectedly large number of predicted genes on this chromosome. The intergenic repetitive fraction of the genome is not well understood and is frequently described as 'junk'. Repetitive sequences are usually removed or separated from other sequences before whole-genome shotgun assembly because they can cause global misassembly. But we know that functional genes are found in repetitive sequences and that transposable elements embedded in the repetitive sequences can restructure genomes, can control gene action and are likely to be involved in generating some of the allelic variation that has been selected in plants.

In addition, high-quality finished sequence provides the only real opportunity to study gene regulation, because most of the essential regulatory sequences fall outside the transcribed regions and our analysis of a restricted region of the genome showed that 43% of the genes predicted from whole-genome shotgun sequence methods were incomplete. Our results and those from the sequencing of rice chromosome 4 (ref. 18) show clearly the importance of the finished sequence. The IRGSP has an immediate goal of sequencing the rice genome to a minimum standard of the high-throughput genomic sequence (HTG) phase 2 level by the end of 2002 and is committed to a long-term goal of obtaining finished high-quality sequence for the whole genome. □

Methods

Chromosome sequencing

We sequenced the whole chromosome 1 of *Oryza sativa* ssp. *japonica*, variety Nipponbare, from 390 overlapping PAC/BAC clones. Initially, we constructed a sequence-ready physical map using the RGP *Sau*3AI PAC and *Mbo*I BAC libraries¹⁹. We also used *Hind*III or *Eco*RI BAC libraries constructed by Clemson University Genomics Institute (CUGI), and BAC clones with draft sequence data provided by Monsanto for gap filling in particular. We carried out shotgun sequencing of RGP and CUGI PAC/BAC clones to obtain sequence data with tenfold overlap. For Monsanto BAC clones²⁰, we complemented the available draft sequence (fivefold redundancy) with an additional fivefold overlap sequence (<http://rgp.dna.affrc.go.jp/genomicdata/seqstrategy/newstrategy.html>).

After the initial assembly of sequence data, stretches of poor or ambiguous quality and apparent gap regions were identified for further sequencing to obtain greater than 99.99% sequence accuracy. But despite extensive efforts to improve the sequence quality and to fill the gaps, 4 of the 390 PAC/BAC clones sequenced are still at phase 1 (GenBank, <http://www.ncbi.nlm.nih.gov/HTGS/>) because the consensus sequence could not be ordered correctly owing to numerous repeats. The remainder comprises 16 phase 2 and 370 phase 3 clones. The nine contigs for chromosome 1 representing the non-overlapping segments of continuous sequence were conjoined by inserting into the gap regions nucleotides that were calculated on the basis of the results of FISH experiments. All of the sequence information of chromosome 1 has been submitted to the DNA Data Bank of Japan (DDBJ, <http://www.ddbj.nig.ac.jp/>) with the accession number BA000010 (Con Division).

Gene prediction and functional classification

We carried out gene prediction using our in-house automated gene prediction system RiceGAAS¹³. The algorithm for gene domain prediction in RiceGAAS was designed by combining several prediction programs including GENSCAN²¹ for maize, GENSCAN²¹ for *Arabidopsis*, RiceHMM (<http://rgp.dna.affrc.go.jp/RiceHMM/index.html>) and the exon-finding program MZEF (<http://argon.cshl.org/genefinder/>), with homology search results from BLASTN and BLASTX (<http://www.ncbi.nlm.nih.gov/BLAST/>). These results were merged and integrated for gene prediction. Domain search was done using InterPro (<http://www.ebi.ac.uk/interpro/scan.html>), and repeats were identified using RepeatMasker (<http://ftp.genome.washington.edu/cgi-bin/RepeatMasker>). The predicted proteins were used to query the nonredundant protein database using BLASTP and categorized according to functional categories defined for *Arabidopsis* by MIPS (http://mips.gsf.de/cgi-bin/proj/thal/filter_funcat.pl?all) with a threshold probability value of 10⁻²⁰.

Received 4 April; accepted 19 September 2002; doi:10.1038/nature01184.

1. Harushima, Y. *et al.* A high-density rice genetic linkage map with 2275 markers using a single F₂ population. *Genetics* **148**, 479–494 (1998).
2. Wu, J. *et al.* A comprehensive rice transcript map containing 6591 expressed sequence tag sites. *Plant Cell* **14**, 525–535 (2002).
3. Yamamoto, K. & Sasaki, T. Large-scale EST sequencing in rice. *Plant Mol. Biol.* **35**, 135–144 (1997).
4. Sasaki, T. & Burr, B. International rice genome sequencing project: the effort to completely sequence the rice genome. *Curr. Opin. Plant Biol.* **3**, 138–141 (2000).
5. Yu, J. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**, 79–91 (2002).
6. Goff, S. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92–100 (2002).
7. Cheng, Z. *et al.* Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. *Plant Cell* **14**, 1691–1704 (2002).
8. Dong, F. *et al.* Rice (*Oryza sativa*) centromeric regions consist of complex DNA. *Proc. Natl Acad. Sci. USA* **95**, 8135–8140 (1998).
9. Richards, E. J. & Ausubel, F. M. Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* **53**, 127–136 (1988).
10. The Arabidopsis Genome Initiative Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–820 (2000).
11. Cheng, Z. *et al.* Toward a cytological characterization of the rice genome. *Genome Res.* **11**, 2133–2141 (2001).
12. Barakat, A., Carrels, N. & Bernardi, G. The distribution of genes in the genomes of Gramineae. *Proc. Natl Acad. Sci. USA* **94**, 6857–6861 (1997).
13. Sakata, K. *et al.* RiceGAAS: an automated annotation system and database for rice genome sequence. *Nucleic Acids Res.* **30**, 98–102 (2002).
14. Hiratsuka, J. *et al.* The complete sequence of rice (*Oryza sativa*) chloroplast genome: Intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. Gen. Genet.* **217**, 185–194 (1989).
15. Lowe, T. & Eddy, S. R. tRNA-Scan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
16. Wessler, S. R., Bureau, T. E. & White, S. E. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Curr. Opin. Genet. Dev.* **5**, 814–821 (1995).
17. Sasaki, A. *et al.* A mutant gibberellin-synthesis gene in rice. *Nature* **416**, 701–702 (2002).
18. Feng, Q. *et al.* Sequence and analysis of rice chromosome 4. *Nature* **420**, 316–320 (2002).
19. Baba, T. *et al.* Construction and characterization of rice genome libraries: PAC library of *japonica* variety, Nipponbare, and BAC library of *indica* variety, Kasalath. *Bull. Natl. Inst. Agrobiol. Resour. (Japan)* **14**, 41–52 (2000).
20. Barry, G. The use of the Monsanto draft rice genome sequence in research. *Plant Physiol.* **125**, 1164–1165 (2001).
21. Burge, C. & Karlin, S. Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**, 78–94 (1997).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements We thank Monsanto for the BAC contig information, BAC clones and their sequence data; R. Wing of Clemson University Genomics Institute and Novartis for the rice Nipponbare BAC library and its fingerprint data, respectively; M. Hattori for technical assistance; B. Burr and F. Burr for critically reading the manuscript; T. Slezak for comments; and K. Eguchi for encouragement.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T. Sasaki (e-mail: tsasaki@nias.affrc.go.jp).

Sequence and analysis of rice chromosome 4

Qi Feng^{††}, Yujun Zhang^{††}, Pei Hao^{††}, Shengyue Wang^{†‡}, Gang Fu[‡], Yucheng Huang^{*}, Ying Li^{*}, Jingjie Zhu^{*}, Yilei Liu^{*}, Xin Hu^{*}, Peixin Jia^{*}, Yu Zhang^{*}, Qiang Zhao^{*}, Kai Ying^{*}, Shuliang Yu^{*}, Yesheng Tang^{*}, Qijun Weng^{*}, Lei Zhang^{*}, Ying Lu^{*}, Jie Mu^{*}, Yiqi Lu^{*}, Lei S. Zhang^{*}, Zhen Yu^{*}, Danlin Fan^{*}, Xiaohui Liu^{*}, Tingting Lu^{*}, Can Li^{*}, Yongrui Wu^{*}, Tongguo Sun^{*}, Haiyan Lei^{*}, Tao Li^{*}, Hao Hu^{*}, Jianping Guan^{*}, Mei Wu^{*}, Runquan Zhang^{*}, Bo Zhou^{*}, Zehua Chen^{*}, Ling Chen^{*}, Zhaoqing Jin^{*}, Rong Wang^{*}, Haifeng Yin[‡], Zhen Cai[‡], Shuangxi Ren[‡], Gang Lv[‡], Wenyi Gu[‡], Genfeng Zhu[‡], Yuefeng Tu[‡], Jia Jia[‡], Yi Zhang[‡], Jie Chen[‡], Hui Kang[‡], Xiaoyun Chen[‡], Chunyan Shao[‡], Yun Sun[‡], Qiuping Hu[‡], Xianglin Zhang[‡], Wei Zhang[‡], Lijun Wang[‡], Chunwei Ding[‡], Haihui Sheng[‡], Jingli Gu[‡], Shuting Chen[‡], Lin Ni[‡], Fenghua Zhu[‡], Wei Chen[§], Lefu Lan[§], Ying Lai[§], Zhukuan Cheng^{¶¶}, Minghong Gu^{¶¶}, Jiming Jiang^{¶¶}, Jiayang Li[§], Guofan Hong^{*}, Yongbiao Xue[§] & Bin Han^{*}

^{*} National Center for Gene Research, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 500 Caobao Road, Shanghai 200233, China

[‡] Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Shanghai 201203, China

[§] Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Datun Road, Andingmenwai, Beijing 100101, China

[¶] Yangzhou University, 27 Wenhua Road, Yangzhou, Jiangsu 225009, China

^{¶¶} Department of Horticulture, University of Wisconsin, Madison, Wisconsin 53706, USA

^{††} These authors contributed equally to this work

Rice is the principal food for over half of the population of the world. With its genome size of 430 megabase pairs (Mb), the cultivated rice species *Oryza sativa* is a model plant for genome research¹. Here we report the sequence analysis of chromosome 4 of *O. sativa*, one of the first two rice chromosomes to be sequenced completely². The finished sequence spans 34.6 Mb and represents 97.3% of the chromosome. In addition, we report the longest known sequence for a plant centromere, a completely sequenced contig of 1.16 Mb corresponding to the centromeric region of chromosome 4. We predict 4,658 protein coding genes and 70 transfer RNA genes. A total of 1,681 predicted genes match available unique rice expressed sequence tags. Transposable elements have a pronounced bias towards the euchromatic regions, indicating a close correlation of their distributions to genes along the chromosome. Comparative genome analysis between cultivated rice subspecies shows that there is an overall syntenic relationship between the chromosomes and divergence at the level of single-nucleotide polymorphisms and insertions and deletions. By contrast, there is little conservation in gene order between rice and *Arabidopsis*.

The rice genome has been well mapped both genetically and physically^{3–5} and has a syntenic relationship with other cereals⁶. *Arabidopsis thaliana* (*Arabidopsis*), a member of the Brassica family of dicotyledonous (dicot) plants, has become an important model flowering plant for studying many aspects of plant biology⁷. The completion of the *Arabidopsis* genome^{8–10} has afforded an unprecedented opportunity for systematic studies of plant gene function. Equally, the complete rice genome sequence will provide a catalogue of genes that may be important for improving not only rice but also other cereals, as functionally important sequences are conserved and may be identified by their similarity¹¹. The International Rice Genome Sequencing Project (IRGSP) has adopted the clone-by-clone approach for obtaining a finished rice genome sequence, because it is modular, allows efficient gap filling, avoids problems arising from distant repetitive sequences and results in the early completion of larger contiguous segments of a genome. Here we describe the completed sequence of rice *O. sativa* ssp. *japonica* cv.

Nipponbare chromosome 4 to an accuracy of greater than 99.99%, together with its analysis.

We constructed a comprehensive clone-based physical map of the Nipponbare chromosome 4 through an integrated approach (ref. 12 and Fig. 1). We identified a tiling path that consisted of 287 bacterial artificial chromosomes (BACs) and two phage (P1)-derived artificial chromosomes (PACs; Supplementary Fig. 1). Each clone was sequenced by a random shotgun approach on both strands of small-insert subclones to achieve a tenfold coverage. Sequence assemblies were verified by comparing *NotI* and *HindIII* restriction profiles predicted from the sequence against experimentally determined restriction profiles.

The collinearity of assembled sequences with the chromosome was confirmed further by integrating 176 sequenced genetic and 402 expressed sequence tag (EST) markers (Supplementary Table 1). Sequence overlaps of adjacent clones were used to assess potential sequence differences between them, which resolved the sequence with less than 1 mismatch per 14,000 base pairs (bp); these mismatches were corrected further by manual checking or by additional sequencing reactions. We estimate the overall accuracy of our finished sequence to exceed 99.99%.

In total, we finished 34,689,786 bp of sequence in eight sequence contigs (Fig. 1). The pachytene spread fluorescence *in situ* hybridization (FISH) and fibre FISH analyses indicated that the sizes of gaps 1–7 are about 125.91 kb, 112.84 kb, 47.82 kb, 110.89 kb, 73.57 kb, 101.06 kb and 301.76 kb, respectively, which leaves 885.49 kb of unsequenced regions (Supplementary Fig. 2 and Supplementary Table 2). The total length of chromosome 4 was therefore calculated to be 35.6 Mb, which is 1.3 Mb shorter than previous estimates of 36.8 Mb (refs 6, 13). Thus, 97.3% (34.6/35.6 Mb) of the chromosome including the centromeric region has been sequenced and

represents a practically finished chromosome 4.

The distribution of various repetitive DNA elements along chromosome 4 is shown in Fig. 2. In addition, details of the content of repetitive sequences in rice chromosome 4 are given in Supplementary Table 3. Overall, the repeats are mainly located in the heterochromatic regions including the centromere. But the miniature inverted-repeats transposable elements (MITEs) have a pronounced bias towards the euchromatic regions along the long arm south (Fig. 2); these are the most frequently occurring class of repetitive elements and, with a total of 7,666, account for 49% of the total number of repetitive DNA elements. No MITEs have been found in exons and they are rare in introns. Other types of repeat seem to be dispersed in the rest of the chromosome, with no obvious clustering. This is similar to the Arabidopsis genome¹⁴, but in contrast to what has been observed in maize, where nested retroelements are the main component of the intergenic regions. Overall, according to an estimation based on current information in the rice RepeatDatabase, the total length of the repetitive sequence is 6.3 Mb, accounting for 18.2% of chromosome 4 excluding the unsequenced regions. The G + C content of chromosome 4 is 44.16%. The heterochromatic region shows a very low recombination frequency with a ratio of physical distance to genetic distance of 1.57 cM Mb⁻¹; however, the euchromatic region has a value of 4.79 cM Mb⁻¹.

We used gene prediction programs and homology searches against all available DNA and protein sequences, in particular rice

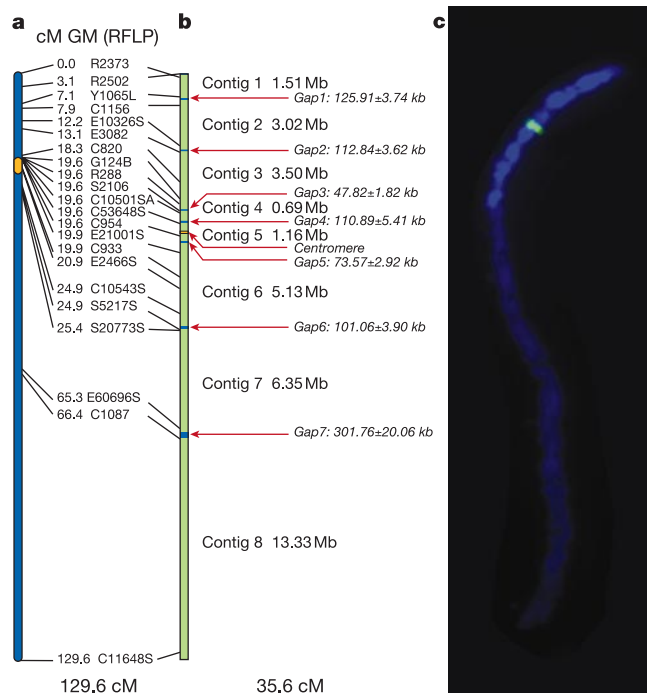


Figure 1 Maps of rice chromosome 4. **a**, Genetic map showing the positions of the centromere (yellow) and genetic markers. **b**, Physical map showing the locations of eight sequenced contigs (green) numbered from the top, seven remaining gaps (blue) indicated by red arrows, and the centromere (yellow) on the chromosome. The sizes of the unsequenced gaps measured by fibre FISH analysis are given in kilobases. The whole chromosome is about 35.53 Mb. **c**, A 4',6'-diamidino-2-phenylindole dihydrochloride (DAPI)-stained pachytene spread of rice chromosome 4 showing heterochromatic (bright blue) and euchromatic (light blue) regions. The centromere (green) was detected by FISH with a centromere-specific probe (pRCS2 clone).

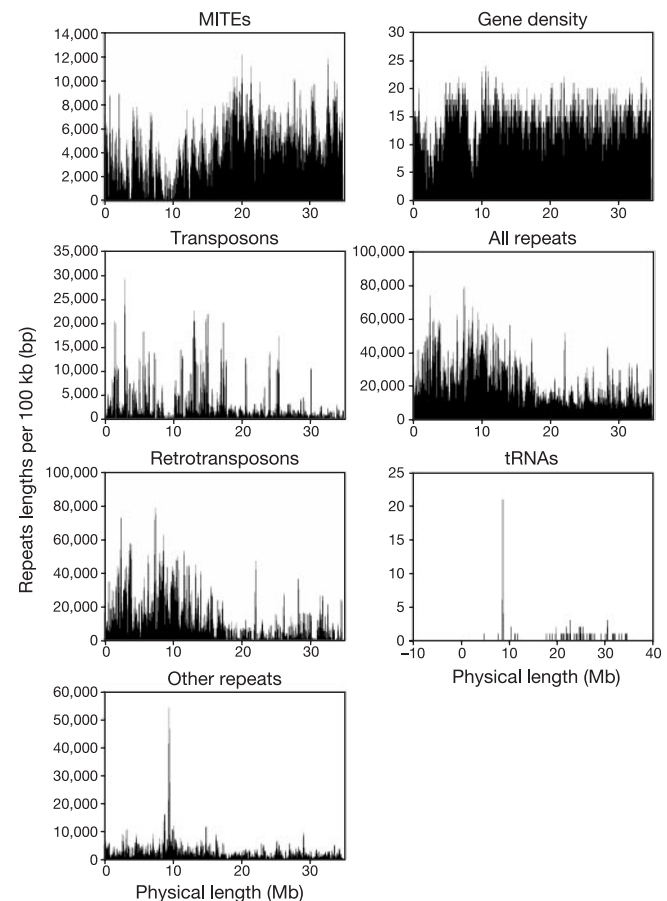


Figure 2 Distribution of various repeats and features along chromosome 4. The chromosome is represented on the x axis and is numbered from base pair 1 of the north telomere to base pair 34,689,786 bp of the south telomere by joining up the physical gaps. The y axis shows the various repeats plotted relative to the nucleotide position on the chromosome. The densities of the repeats and tRNA were obtained by analysing the sequence every 100 kb using a 10-kb sliding window.

Table 1 General structural features of rice chromosome 4

Statistics of annotated chromosome 4	
Total genes	4,658
Average gene size (bp)	
Gene density (bp per gene)	7,441
Average exon per gene	4.4
Average exon size (bp)	340
Average intron size (bp)	376
Average BAC GC (%)	44.16
Average gene GC (%)	47.22
Average intergenic GC (%)	42.30
Average exon GC (%)	53.0
Average intron GC (%)	39.0

EST sequences including 115,000 reads from *indica* Guangluai 4 generated by ourselves, to annotate the rice genomic sequences by *ab initio* methods. This analysis relied much more heavily on the results of a manual curation. There are a total of 4,658 coding genes predicted on chromosome 4. The average gene density is 7,441 bp per gene, which is lower than in *Arabidopsis* (4.5 kb per gene). We estimate that the total number of rice genes is about 57,000, with the assumption that the rest of the genome has a similar gene density to that of chromosome 4. On average, each gene contains 4.4 exons and is about 2,779 bp. The largest gene contains 32 exons and encodes a protein of 1,483 amino acids with 69% similarity to the polyadenylation protein of *Arabidopsis*¹⁵. Seventy tRNA and four small nucleolar RNA genes were found. Twenty-four of the tRNA genes are chloroplast-derived tRNA genes located in a cluster near centromere (Supplementary Table 4). We identified 1,681 genes to have unique rice EST matches. As only a limited number of ESTs were available, we classified 2,073 of the predicted genes as hypothetical genes. The remaining 2,585 predicted genes have been classified as 'known', 'novel' or 'putative' because they have EST matches, protein matches or any detectable domain signatures. The gene map of chromosome 4 is shown in Supplementary Fig. 1 together with its other features. The statistics of the annotated chromosome 4 and its general features are given in Table 1.

To classify the function of the predicted genes, we used BLASTP (basic local alignment search tool) to search for homologous proteins in protein databases, as well as InterPro to search for domain signatures. It was difficult to assign a function to most of the predicted coding proteins, and these are therefore classified as

'others'. Clearly, the proteins involved in cellular metabolism (555 proteins, 11.92%), transcription (259, 5.56%) and signal transduction (256, 5.51%) typically represent a high proportion of the protein set (Table 2). Notably, 161 predicted proteins containing a disease resistance protein signature and other associated domains were identified to be involved in plant defense. Among these, a bacterial blight-resistance gene *Xa1* cluster consisting of six members was identified. As it has been reported that *Xa1* gene is a single-copy gene in the rice genome¹⁶, it will be interesting to test whether each member of the *Xa1* gene cluster has a role in recognizing different races of the rice bacterial pathogen *Xanthomonas oryzae* pv. *oryzae*. In total, 1,004 genes are in gene families with more than two members on the chromosome, indicating a high proportion (at least 21.5%) of gene families in the rice genome (Supplementary Fig. 3).

The sequences of ten minimally tiled BACs of the 1.16-Mb contig derived from the centromeric region have been determined completely (Supplementary Fig. 4). Sequences in a BAC clone OSJNBb0062N22 located 0.4 Mb from the top end of the contig have a significant similarity to a satellite repeat (pRCS2) associated with rice centromeres¹⁷. The physical structure of the core centromeric sequence consists of 341 copies of pRCS2-like repeats with 80–100% similarity. There are 15 genes predicted in this region. Among them, ten seem to be related to retroelements and five are hypothetical. Thirteen have unique rice EST matches. Because the centromere is apparently not subject to recombination, these genes will be useful to investigate chromosome evolution at the molecular level. The dispersed repeats consist predominantly of long terminal repeats and simple sequence repeats, which are all found frequently in the heterochromatin (Supplementary Fig. 5). The boundaries of the centromere could be defined further by cytogenetic and functional analysis. Nevertheless, the centromere-related sequence of chromosome 4 provides both a structural basis for studying centromere function in plants and an opportunity to construct rice or plant artificial chromosomes that can be potentially used in the cloning and transformation of large DNA fragments.

The cultivated rice species *O. sativa* is classified into two main subspecies, *indica* and *japonica*¹⁸, that represent most of the rice crops grown in the world¹⁹. We carried out sequence alignments between 2.3 Mb of three contiguous segments of chromosome 4 from *indica* Guangluai 4 and its collinear sequences from *japonica* Nipponbare, which revealed extensive sequence collinearity including gene

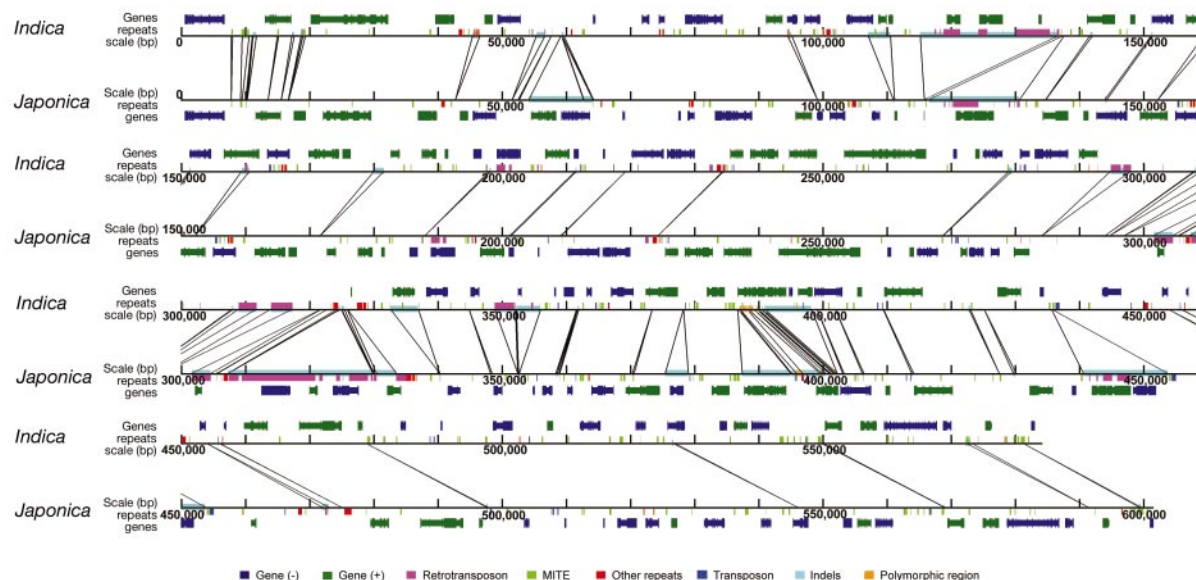


Figure 3 Comparative sequence analyses between rice subspecies *indica* GLA4 and *japonica* Nipponbare. The *indica* (top) and *japonica* (bottom) sequences were aligned over contiguous segments of 583 kb (*indica*) and 602 kb (*japonica*). Insertions and deletions

are indicated in light blue. Genes on the two strands are indicated in green and blue, respectively. Polymorphic regions are indicated in orange, and other types of repeats are indicated in different colours.

Table 2 Functional classification of rice chromosome 4

Cellular metabolism	555	11.92%
Transcription	259	5.56%
Signalling	256	5.51%
Transport	198	4.14%
Plant defence	161	3.47%
Protein fate	85	1.84%
Growth	81	1.76%
Other	3,063	65.76%
Total	4,658	

contents and orders between them. But deviations from the collinearity are frequent with insertions or deletions (Indels; Fig. 3). Chromosome 4 of *japonica* is probably larger than that of *indica* because of an expansion by insertions of transposable elements. If this were also true for other chromosomes, it would mean that the *indica* genome is smaller than that of the *japonica*. The insertions occur not only in the intergenic regions, but also in some of the coding regions. We identified 9,056 single-nucleotide polymorphisms (SNPs) in 2.3 Mb of sequence, indicating an average frequency of 1 SNP per 268 bp. There are 63 Indels for the *indica* sequence and 138 Indels for the *japonica* sequence. The results are summarized in Supplementary Table 5. These findings may reflect an overall conservation of the two genomes. Because the *indica* and *japonica* rice species had a polyphyletic origin from a common wild annual rice species *O. nivara*¹⁹, some of the differences identified from the two chromosomes may simply have resulted from the differences of the two varieties used. We need more genomic data from other rice varieties of *indica* and *japonica* to identify the evolutionary relationships of the cultivated rice species.

The working drafts of the rice *indica* variety 93-11 genome and the *japonica* Nipponbare have been completed^{20,21}. We identified 7,423 contigs with a length of 26,142,459 bp from a total of 127,550 contigs of the 93-11 draft sequences on chromosome 4 by using a similarity of >90% as a match cutoff. These contigs matched 7,551 regions on the *japonica* Nipponbare chromosome 4, indicating 128 possible insertions on this chromosome, and covered 75.42% of the chromosome (Supplementary Fig. 6). We estimated that about 2,225 predicted genes (47.77%) are present in full in the 93-11 draft sequences, 1,788 genes (38.38%) are partially identified, and 645 genes (13.85%) are not found from the located draft sequences. All of these identified free contigs of the 93-11 draft sequences might serve as a scaffold for comparative genome analysis of *indica* and *japonica*.

Both rice and Arabidopsis are model angiosperm plants, but they are distantly related species. We compared the proteins predicted on rice chromosome 4 with a complete set of proteins of *Arabidopsis thaliana* using the dps program (dna against protein database search) of the AAT package with stringency value of parameter $f > 200$ (ref. 22). We identified 2,040 Arabidopsis proteins with significant homology in amino acid sequence. The distribution of rice genes that encode protein products related to the Arabidopsis proteins has a pronounced bias toward the euchromatic region of rice chromosome 4 (Supplementary Fig. 7). It is evident that the 130–200 million years that separate the two main angiosperm groups²³ have eroded conservation of gene order to the point where the comparative approach is no longer a useful tool with predictive power, consistent with previous observations that suggest that collinearity between dicot and monocot species seems to be limited to a few very small regions. The remaining 2,618 proteins that have no significant homology to Arabidopsis proteins may represent rice- or monocot-specific proteins. In this set, 14 protein-coding genes matched maize unique ESTs. Two of these proteins have known functions in maize (a phytase for hydrolysis of phytin²⁴) and in coconut (a lysophosphatidic acid acyltransferase for accepting medium-chain-length substrates²⁵), respectively. Of 1,794 maize marker sequences, 114 markers were mapped on the rice chromo-

some 4 (Supplementary Fig. 8). Macrocollinearities between rice chromosome 4 and maize chromosome 2 and 10 are observed, but microcollinearity between the two genomes was not apparent.

The sequence of rice chromosome 4 provides us with further insight into the structure and organization of eukaryotic chromosomes. Functional analysis of the 65% of genes on rice chromosome 4 that are predicted but have unknown function clearly represents a considerable challenge. The overall breakdown of genetic synteny between monocots and dicots, seen through the comparison of rice chromosome 4 and Arabidopsis, suggests that extensive genome reshuffling has occurred since their divergence. The comparison between rice subspecies, as well as varieties, shows that the principal sequence differences include SNPs and Indels, which might be useful for dissecting the molecular basis of the phenotypic differences that are often associated with different subspecific or varietal lines. □

Methods

Chromosome sequencing

Four genomic libraries made from the rice *O. sativa* ssp. *japonica* cv. Nipponbare were used to sequence chromosome 4: two BAC libraries of OSJNBa and OSJNBb were provided by the Clemson University Genomics Institute (CUGI); a PAC library was provided by the Rice Genome Project (RGP); and some BACs were provided by Monsanto. We purified the selected BACs and PACs by a caesium chloride gradient. For the shotgun approach, sheared BAC or PAC DNA (3 kb) was ligated into a pBluescript vector and transformed into *Escherichia coli* strain DH5 α . The shotgun subclones with inserts of 3 kb were sequenced from both ends by the dideoxy chain termination method by using either BigDye Terminator Cycle Sequencing V2.0 Ready Reaction (Applied Biosystems) or Dyanamic ET Dye Terminator kit (MegaBACE; Amersham Pharmacia). Most of the reactions were analysed on MegaBACE 1000 Capillary Sequencing machines (Amersham Pharmacia). ABI377 and ABI3700 sequencers (Applied Biosystems) were also used in this study. The subclones were sequenced to generate between eight- and tenfold coverage of each BAC or PAC clone. Individual BACs or PACs were assembled from the shotgun sequences using the PHRED and PHRAP²⁶ programs. For sequence finishing, we used the GAP4 program of the Staden Packages to help edit and select the additional sequencing reactions for gap closure and problem solving. The *Consed* program²⁷ was also used for sequence finishing. Sequence gaps were closed by using various dye-labelled terminator chemistries or by a combination approach of primer walking, small- or big-insert library constructing, and polymerase chain reaction with oligonucleotides.

Gene prediction and sequence alignments

We analysed the DNA sequence of each BAC for protein-coding genes using gene-prediction software and alignments with the sequences from the EST and nonredundant protein databases. For initial gene and exon predictions, we used FGENESH (<http://www.softberry.com/>), Genscan²⁸ and GeneMark.HMM (<http://dixie.biology.gatech.edu/Genemark/eukhmm.cgi>). We aligned sequences in the EST and nonredundant protein databases to the BAC sequences using BLAST²⁹ and AAT (<http://genome.cs.mtu.edu/aat.html>) sequence database search and alignment software. We identified repeats using RepeatMasker2 (<http://ftp.genome.washington.edu/cgi-bin/RepeatMasker>) and used diffseq (<http://www.hgmp.mrc.ac.uk/Software/EMBOSS/Apps/diffseq.html>) to find SNPs.

We used tRNA scan-SE³⁰ and manual inspection to identify cytoplasmic tRNAs and organelle-derived tRNA.

Received 4 April; accepted 16 September 2002; doi:10.1038/nature01183.

1. Sasaki, T. & Burr, B. International rice genome sequencing project: the effort to completely sequence the rice genome. *Curr. Opin. Plant Biol.* **3**, 138–141 (2000).
2. Sasaki, T. *et al.* The genome sequence and structure of rice chromosome 1. *Nature* **420**, 312–316 (2002).
3. Harushima, Y. *et al.* A high-density rice genetic linkage map with 2275 markers using a single F₂ population. *Genetics* **148**, 479–494 (1998).
4. Wu, J. *et al.* A comprehensive rice transcript map containing 6591 expressed sequence tag sites. *Plant Cell* **14**, 525–535 (2002).
5. Chen, M. *et al.* An integrated physical and genetic map of the rice genome. *Plant Cell* **14**, 537–545 (2002).
6. Gale, M. D. & Devos, K. M. Comparative genetics in the grasses. *Proc. Natl Acad. Sci. USA* **95**, 1971–1974 (1998).
7. Meinke, D. W., Cherry, J. M., Dean, C. D., Rounsley, S. & Koornneef, M. *Arabidopsis thaliana*: a model plant for genome analysis. *Science* **282**, 662–682 (1998).
8. The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815 (2000).
9. Lin, X. *et al.* Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* **402**, 761–768 (1999).
10. Mayer, K. *et al.* Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*. *Nature* **402**, 769–777 (1999).
11. Messing, J. & Llaça, V. Importance of anchor genomes for any plant genome project. *Proc. Natl Acad. Sci. USA* **95**, 2017–2020 (1998).
12. Zhao, Q. *et al.* A fine physical map of the rice chromosome 4. *Genome Res.* **12**, 817–823 (2002).
13. Saji, S. *et al.* A physical map with yeast artificial chromosome (YAC) clones covering 63% of the 12 rice chromosomes. *Genome* **44**, 32–37 (2001).
14. Copenhaver, G. *et al.* Genetic definition and sequence analysis of *Arabidopsis* centromeres. *Science* **286**, 2468–2474 (1999).

15. Jenny, A. & Keller, W. Cloning of cDNAs encoding the 160kDa subunit of the bovine cleavage and polyadenylation specificity factor. *Nucleic Acids Res.* **23**, 2629–2635 (1995).
16. Yoshimura, S. *et al.* Expression of *Xa1*, a bacterial blight-resistance gene in rice, is induced by bacterial inoculation. *Proc. Natl Acad. Sci. USA* **95**, 1663–1668 (1998).
17. Dong, F. *et al.* Rice (*Oryza sativa*) centromeric regions consist of complex DNA. *Proc. Natl Acad. Sci. USA* **95**, 8135–8140 (1998).
18. Oka, H. I. in *Rice Biotechnology* (eds Khush, G. S. & Toennissen, G. H.) 55–80 (CAB International, Oxon, 1991).
19. Khush, G. S. Origin, dispersal, cultivation and variation of rice. *Plant Mol. Biol.* **35**, 25–34 (1997).
20. Yu, J. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**, 79–92 (2002).
21. Goff, S. A. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92–100 (2002).
22. Huang, X., Adams, M. D., Zhou, H. & Kerlavage, A. R. A tool for analyzing and annotating genomic sequences. *Genomics* **46**, 37–45 (1997).
23. Paterson, A. H. *et al.* Toward a unified genetic map of higher plants, transcending the monocot–dicot divergence. *Nature Genetics* **14**, 380–382 (1996).
24. Maugeness, S., Martinez, I., Godin, B., Perez, P. & Lescure, A. M. Structure of two maize phytase genes and their spatio-temporal expression during seedling development. *Plant Mol. Biol.* **39**, 503–514 (1999).
25. Knutzon, D. S. *et al.* Cloning of a coconut endosperm cDNA encoding a 1-acyl-sn-glycerol-3-phosphate acyltransferase that accepts medium-chain-length substrates. *Plant Physiol.* **109**, 999–1006 (1995).
26. Ewing, B. & Green, P. Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res.* **8**, 186–194 (1998).
27. Gordon, D., Abajian, C. & Green, P. Consed. A graphical tool for sequence finishing. *Genome Res.* **8**, 195–202 (1998).
28. Burge, C. & Karlin, S. Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**, 78–94 (1997).
29. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
30. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).

Supplementary Information accompanies the paper on Nature's website
(<http://www.nature.com/nature>).

Acknowledgements We thank T. Sasaki and the RGP for rice genetic and EST markers and a PAC genomic library of rice Nipponbare; R. Wing and the CUGI for providing BAC libraries of the Nipponbare variety; Monsanto for the rice working-draft sequence data; G. Barry and J. Liu for help; R. Buell and Q. Yuan for help with the annotation and analysis of chromosome 4 sequences; X. Huang and Z. Ning for help with using the AAT and the *ssaha* programs, respectively; members of the National Centre for Gene Research for assistance; Z. Xu, Z. Chen, G. Wang, Q. Ma and Q. Zhang for support; and X. Lin, X. Deng, Y. Li, L. Zhou, N. Zheng, X. Liu and members of the IRGSP for discussion. This work was supported by grants from the Ministry of Science and Technology of the People's Republic of China, Chinese Academy of Sciences, and the Shanghai Municipal Commission of Science and Technology.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to B.H.
(e-mail: bhan@ncgr.ac.cn).

Multiplicative computation in a visual neuron sensitive to looming

Fabrizio Gabbiani*†, Holger G. Krapp*‡, Christof Koch* & Gilles Laurent*

* Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

† Division of Neuroscience, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA

‡ Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

Multiplicative operations are important in sensory processing^{1–5}, but their biophysical implementation remains largely unknown^{6–10}. We investigated an identified neuron (the lobula giant movement detector, LGMD, of locusts) whose output firing rate in response to looming visual stimuli has been described by two models, one of which involves a multiplication. In this model, the LGMD multiplies postsynaptically two inputs (one excitatory, one inhibitory) that converge onto its dendritic tree^{11,12}; in the other model, inhibition is presynaptic to the

LGMD^{13,14}. By using selective activation and inactivation of pre- and postsynaptic inhibition, we show that postsynaptic inhibition has a predominant role, suggesting that multiplication is implemented within the neuron itself. Our pharmacological experiments and measurements of firing rate versus membrane potential also reveal that sodium channels act both to advance the response of the LGMD in time and to map membrane potential to firing rate in a nearly exponential manner. These results are consistent with an implementation of multiplication based on dendritic subtraction of two converging inputs encoded logarithmically, followed by exponentiation through active membrane conductances.

Several insect behaviours rely on tracking motion in depth: landing¹⁵, hovering flight¹⁶ and collision avoidance^{17–19}. These behaviours probably depend on different neural computations as animals actively move towards a target or, conversely, experience the approach of a moving threat. The LGMD (Fig. 1a) is an identified neuron located in the third visual neuropile of the locust optic lobe (lobula), and is part of a circuit thought to be involved in the generation of escape behaviours (Fig. 1b). It responds vigorously to solid objects approaching on a collision course with the animal^{17,18} (Fig. 1b, c). The LGMD fires throughout object approach with a rate that increases, peaks, and decreases as collision becomes imminent^{11,12}. Responses are typically brisker for small or fast-moving objects, leading to higher peak firing rates (f_{peak} , Fig. 1d, top). But the timing of the peak firing rate is independent of the object's approach speed or size, and always follows with a fixed delay the time at which the object reaches a fixed threshold angular size, θ_{thres} , on the retina¹². This is seen by plotting the timing of the peak relative to collision (t_{peak}) as a function of the ratio $l/|v|$, where l is the object's half-size and $|v|$ its approach speed (Fig. 1b, d, bottom). The trigonometric relationship between l , v and the angular size, $\theta(t)$, subtended by the object during approach causes the points of such a graph to fall on a straight line if t_{peak} occurs with a fixed delay, δ , relative to θ_{thres} (refs 11, 12). The slope, α , of this straight line is related to θ_{thres} (ref. 12) and its y intercept equals δ . Such a linear relation is indeed observed experimentally^{11,12} (Fig. 1d, bottom; $\theta_{\text{thres}} = 17^\circ$, $\delta = 25$ ms). Thus for the LGMD in this animal, the peak firing rate always occurred 25 ms after the object reached 17° in angular size, independent of v or l .

The peak firing rate of the LGMD was not the only parameter related to threshold angular size during object approach: the time, t_{thres} , at which the firing rate reached a given value, here arbitrarily set at 50 spikes s^{-1} (Fig. 1c triangle), also fell on a straight line ($n = 10$) as a function of $l/|v|$ (Fig. 1d, bottom), and therefore anticipated with a fixed delay (here 101 ms, Fig. 1d) the time at which the object's angular size reached a fixed value (10°). This implies that decoding the LGMD's output by detecting either the peak or an arbitrary threshold instantaneous firing rate ($>50 \text{ spikes s}^{-1}$) conveys a reliable indicator of angular size during looming. The delay between angular threshold and peak firing time is independent of body temperature, arousal level, contrast, variations in shape or texture of the approaching object and direction of approach^{12,20}. Thus, angular threshold might be the image-based retinal variable used to trigger escape responses in the face of an impending collision. Indeed, a leg flexion (presumably in preparation for an escape jump) has been shown to follow the peak LGMD firing rate with a fixed delay¹¹. Angular object size is also closely related to obstacle avoidance behaviours in tethered flying locusts¹⁹.

How then does the LGMD compute this angular threshold? The LGMD receives, onto a large dendritic fan, excitatory retinotopic inputs that convey, to a first approximation, the angular velocity, $\dot{\theta}(t)$ of the approaching object²¹ (field A, Fig. 1a and b, green; Supplementary Information). In addition, two dendritic fields arborize in distinct regions of the lobula and receive phasic non-retinotopic, feedforward inhibition related to object size, $\theta(t)$

(refs 11, 22; fields B and C, Fig. 1a and b, red). We therefore investigated phenomenological models based on combinations of angular velocity and size^{11,12} that could describe the responses of the LGMD to approaching objects. The firing rate of the LGMD was fitted with functions of $\dot{\theta}(t)$ and $\theta(t)$ using a fixed number of parameters. Additive combinations of motion-dependent excitation and size-dependent inhibition could not fit experimental firing rate profiles. By contrast, fits based on the multiplication of an excitatory angular velocity term by an inhibitory (negative) exponential of object size, $[\dot{\theta} \exp(-\alpha\theta)]$, accounted well for most of the response time-course over a wide range of $l/|v|$ values in 10 neurons (Fig. 1c, blue line; Supplementary Information equation (1)). Theoretical arguments¹² show that only this multiplicative combination of excitation and inhibition predicts the linear relationship between peak firing rate and $l/|v|$ observed experimentally (Fig. 1d, bottom). These results suggest that the LGMD multiplies its two types of synaptic inputs related to motion and size of an approaching object.

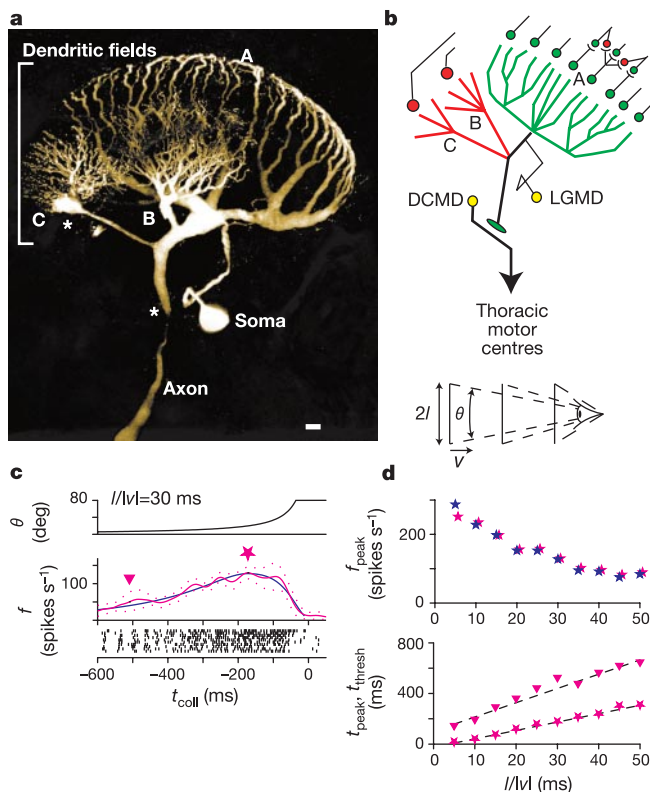


Figure 1 Properties of a neuronal circuit involved in locust escape behaviour and obstacle avoidance. **a**, The LGMD's large dendritic fan (A) arborizes in the lobula. Two additional dendritic arborizations (B, C) receive distinct synaptic inputs (asterisks represent dendritic and axonal recording sites of Figs 3 and 4; scale bar, 20 μ m). **b**, top, fan A receives feedforward excitation (green), while B and C receive feedforward inhibition (red). Lateral inhibition (red) is presynaptic on the excitatory pathway (A). The LGMD synapses onto the DCMD, a neuron that relays spikes in a 1:1 manner to thoracic motor centres. Bottom, diagram illustrating the presentation of approaching squares (half-size, l ; velocity, $v < 0$; ref. 12). $\theta(t)$ depends only on $l/|v|$. **c**, Responses to approaching squares ($l/|v| = 30$ ms). Top, angular size of the stimulus as a function of time relative to collision (t_{coll}). Middle, mean instantaneous firing rate $\pm$ s.d. (magenta line and dots). Fit with multiplicative model in blue (Supplementary Information). Bottom, spike rasters (10 trials). Star, peak instantaneous firing rate. Triangle, threshold firing rate of 50 spikes s^{-1} . **d**, Top, peak instantaneous firing rate as a function of $l/|v|$ (maximal rate: 260 ± 85 spikes s^{-1} , mean $\pm$ s.d., 10 neurons). Blue stars, obtained from model fit as in **c**; mean correlation with $l/|v|$: -0.88 ± 0.19 (10 animals). Bottom, relation between peak (stars) or threshold (triangles) firing times relative to collision and $l/|v|$ is linear.

The excitatory motion-sensitive afferents to the LGMD are also subject to lateral inhibition mediated by a presynaptic network (Fig. 1b) that reduces responses to translating objects^{21,23}. Thus, an alternative model of the LGMD's responses to approaching objects has been proposed, where presynaptic and lateral—rather than postsynaptic and feedforward—inhibition competes with motion-dependent excitation during approach. In this model, feedforward postsynaptic inhibition intervenes only after object motion ceases^{13,14}. Numerical simulations showed that appropriately timed lateral inhibition could indeed play an important role in controlling excitation during object approach¹³. However, because object angular size, $\theta(t)$, depends nonlinearly on time during the approach, it is not clear whether experimental results obtained with translating objects carry over to approaching objects. Lateral inhibition is mediated by local interactions^{21,23} and as the object's angular speed increases during approach, it is expected to have decreasing influence because its speed of activation is limited by propagation time and synaptic delays¹³.

We set out to compare the contributions of feedforward (postsynaptic) and lateral (presynaptic) inhibition in controlling excitation during looming. We designed stimuli that preferentially activate either form of inhibition, and monitored their effect on the LGMD's firing profiles and peak firing time. Figure 2a illustrates one of eight experiments during which an approaching object (top)

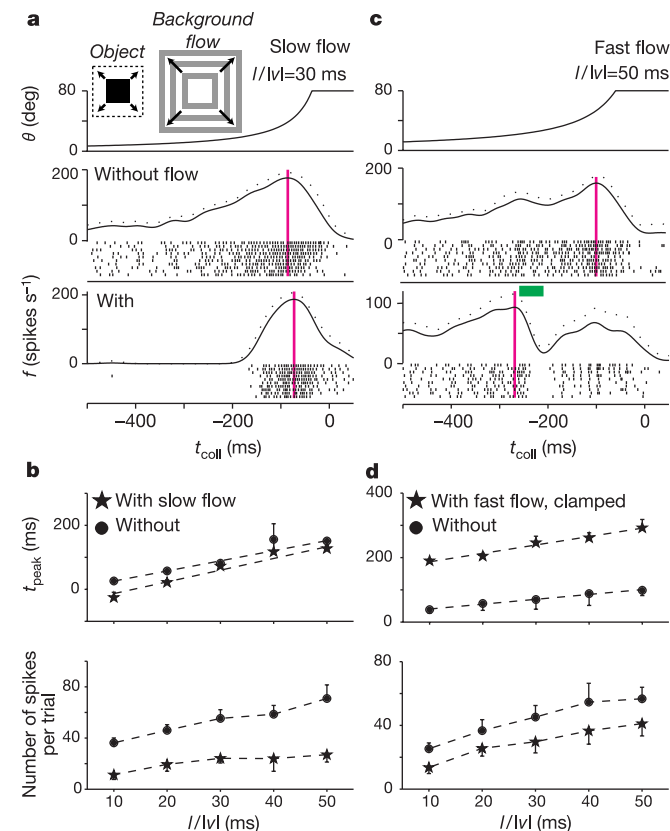


Figure 2 Effect of activating lateral or feedforward inhibition on peak firing time (magenta lines). **a**, Approaching stimulus (top) was presented either alone (middle) or in the presence of a pattern of white/grey stripes (bottom) slowly flowing outwards (inset). **b**, Peak firing time relative to collision with or without background pattern (top) and mean spike number (bottom) for each $l/|v|$ value (mean $\pm$ s.d. 10 repetitions). **c**, Same as **a**, but the background pattern suddenly appeared (green horizontal bar) and moved outwards at a higher speed. **d**, Same as **b** for stimulus presented in **c**. In top panel, stars represent mean peak firing time in trials where it was clamped owing to background pattern motion. Data in **a**, **b** are from a different neuron than those in **c**, **d**.

was either shown alone (top raster) or superimposed on a patterned background (bottom raster) flowing slowly outwards from the centre of the screen (inset). Such slow-flowing stimuli predominantly activate lateral presynaptic inhibition without eliciting inhibitory postsynaptic potentials (IPSPs), that is, without invoking feedforward inhibition (Fig. 3; refs 21, 22). When shown in isolation, these stimuli did not elicit responses in the LGMD, except for 1–2 spikes at the onset of motion. When combined with a looming stimulus, the flowing background caused a significant reduction in the number of spikes fired by the LGMD at all $l/|v|$ values (Fig. 2b, bottom; $P < 0.0005$, rank-sum test). Yet, this reduction occurred exclusively early during the approach: as soon as the object exceeded $23.6^\circ \pm 12.2^\circ$ (8 animals, 5 $l/|v|$ values per animal), the LGMD's responses were virtually unaffected, except for a slight delay in peak firing time (Fig. 2a, b). Neither peak firing rate nor the number of spikes fired in a 100-ms window around the peak were statistically different under both conditions ($P > 0.1$, signed rank-sum test; $n = 8$).

These results contrast sharply with those obtained when the same patterned background appeared abruptly 200 ms before the end of looming (Fig. 2c, green bar) and flowed outwards for 50 ms at a high speed (10 animals; Methods). Such fast-flow stimuli effectively activate postsynaptic inhibition (refs 21, 22; Fig. 2d, bottom; $P < 0.05$, rank-sum test), a result confirmed by intracellular recordings (Fig. 3a). Activity of the LGMD ceased abruptly about 40 ms after fast-flow onset (Fig. 2c, bottom rasters). More

importantly, in 70% of approaches (for 10 animals) the time of peak firing was closely related to the time of fast-flow onset (Fig. 2d, top). Hence, for all values of $l/|v|$ tested, lateral inhibition could not antagonize excitation when the approaching stimulus was closest to collision ($\theta > 23.6^\circ$; Fig. 2a), whereas postsynaptic inhibition was highly effective, even during periods of robust LGMD excitation (Fig. 2c). These results indicate that, as contact nears, postsynaptic inhibition of the LGMD becomes increasingly important in controlling its firing.

We further investigated the time course of inhibition during object approach using local injection of picrotoxin (PCTX), a blocker of chloride channels^{24,25} (Methods; $n = 10$). The drug was applied to dendritic fields B and C of the LGMD. Because lateral inhibition is thought to be mediated by muscarinic acetylcholine receptors in the lobula²³, this manipulation selectively inactivated postsynaptic inhibitory inputs to the LGMD. Figure 3 illustrates LGMD responses to approaching objects before (Fig. 3a) and after (Fig. 3b) PCTX application. Block of postsynaptic inhibition during object approach increased response strength (Fig. 3c, left; $P < 0.008$, rank-sum test; $n = 10$) and peak instantaneous firing rate (191 $\pm$ 60%, mean $\pm$ s.d., $n = 10$). This delayed the peak firing time (Fig. 3c, right; $P < 0.008$, rank-sum test; $n = 10$) and prolonged the responses. To estimate the time at which postsynaptic inhibition started to control excitation significantly, we computed the angle subtended by the stimulus when the mean instantaneous firing rates in control and PCTX started to differ ($P > 0.05$, t -test).

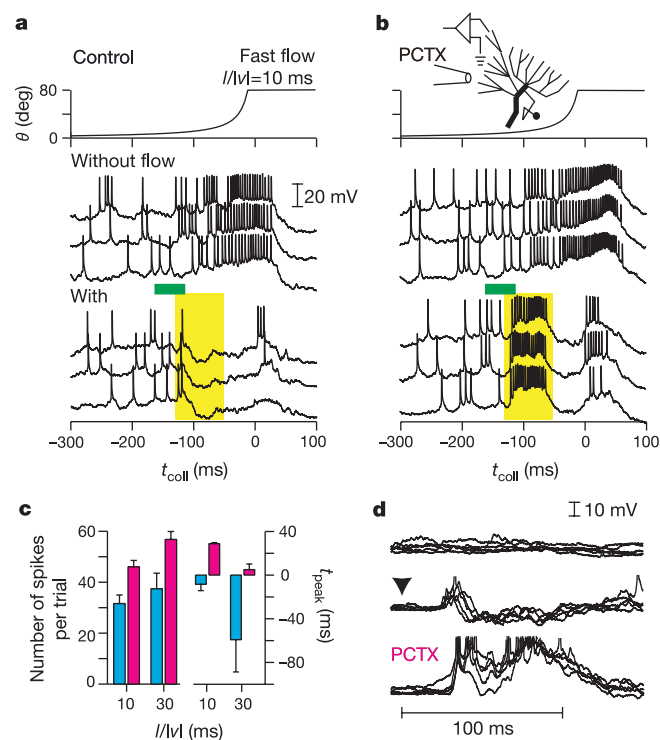


Figure 3 Effect of picrotoxin (PCTX) on the LGMD's responses to approaching stimuli. **a**, Middle 3 intracellular traces illustrate responses to an approaching stimulus (top, dendritic recording; **b**, schematic inset). Bottom 3 traces are responses to the same stimulus with the fast flow (green bar). Note the IPSPs (yellow). **b**, Same protocol after PCTX application (inset). The bottom 3 traces reveal excitation at the time when inhibition was previously present (yellow). **c**, Mean number of spikes (left) and mean peak firing time (right) before (blue) and after (red) PCTX application at two values of $l/|v|$ ($\pm$ s.d., 5 repetitions). **d**, Presentation of the fast flow alone (arrow head, middle traces) elicits EPSPs followed by long-lasting IPSPs. In PCTX, IPSPs are replaced by EPSPs that elicit bursts of spikes (truncated). Top, responses to background alone. Data in **d** are from a different neuron than those in **a–c**.

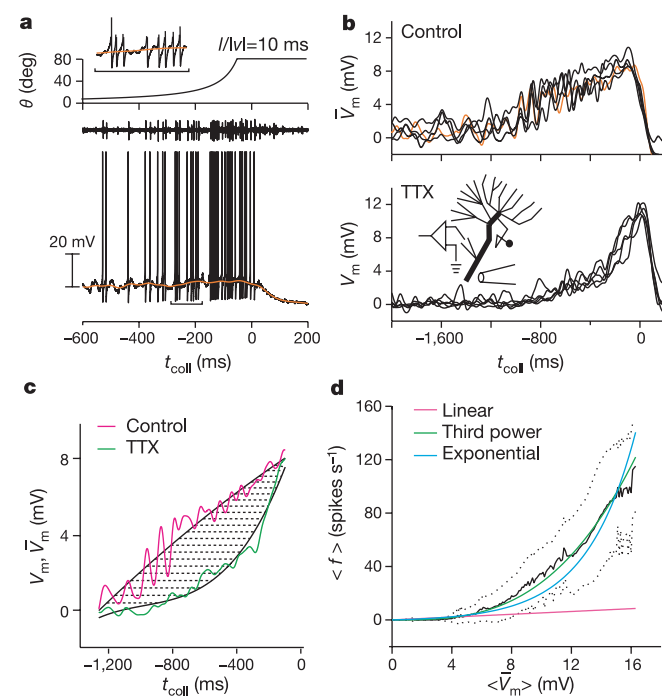


Figure 4 Transformation between membrane potential (V_m) and firing rate at the spike initiation zone. **a**, Approaching stimulus (top), recordings from the DCMD (middle, extracellular) and from the LGMD (bottom, intracellular) close to its spike initiation zone (**b**, inset). Orange trace is membrane potential after median filtering ($\bar{V}_m$). Inset, bracketed portion of V_m and ($\bar{V}_m$) expanded 3 times. **b**, Top panel presents median filtered membrane potential (orange line is same trace as in **a**; 5 repetitions). Bottom 5 traces were recorded after TTX application to the axon (inset). **c**, Mean traces in control and TTX (from **b**) were fitted with a third-order polynomial (black) and used to compute the mean temporal difference (352 ms) in membrane potential over the response rising phase. **d**, Fit of mean instantaneous firing rate, $\langle f \rangle$, as a function of mean, median filtered membrane potential (mean $\pm$ s.d.; solid and dotted black lines) with linear, third-power and exponential models.

The average angle was $23.3^\circ \pm 19^\circ$ (mean $\pm$ s.d., 10 animals and 4 $l/|v|$ protocols per animal), that is, similar to that beyond which lateral inhibition normally loses its influence on the LGMD's responses to an approaching object (Fig. 2a). From then on, the difference between instantaneous firing rate in control and PCTX steadily increased to reach 191% at the peak, consistent with inhibition depending on object size¹¹.

We then presented the approaching object in conjunction with the transient, fast outward flow shown in Fig. 2. Large IPSPs were seen in controls (Fig. 3a, yellow), consistent with previous results (Fig. 2c). Following PCTX injection, a strong excitatory response of duration similar to the presentation of the fast outward flow was revealed (Fig. 3b, yellow). Thus, the fast outward-flowing pattern normally activates concurrent excitation and inhibition, with inhibition rapidly dominating. To confirm this hypothesis we presented the transient fast outward-flow stimulus without approaching object, again in the absence and presence of PCTX (Fig. 3d; $n = 4$). In controls, an excitatory postsynaptic potential (EPSP) arose shortly after the onset of the flow stimulus (arrowhead), before being curtailed by a long-lasting IPSP. After PCTX injection, the same stimulus caused a strong and long-lasting excitation (Fig. 3d). We conclude that postsynaptic inhibition, presumably mediated by GABA_A synapses^{23,24}, potently controls the competing postsynaptic excitation.

Taken together, these results favour a scheme based on the control of excitation by postsynaptic rather than presynaptic inhibition during object approach. Furthermore, they show that the onset of postsynaptic inhibition is tightly matched to the fading of presynaptic inhibition. Under these conditions, a plausible model for the biophysical implementation of the multiplication of angular velocity (θ) and angular size ($\exp(-\alpha\theta)$) components could be by way of a postsynaptic subtraction of their logarithms²⁶, followed by exponentiation^{11,12}. This amounts to implementing $a \times (1/b)$ as $\exp(\log a - \log b)$, a technique commonly used in analogue integrated circuits²⁷. Because our phenomenological fits of the LGMD's responses characterize its firing rate as a function of $\theta \exp(-\alpha\theta)$, the exponentiation of $\log \theta - \alpha\theta$ could occur as the membrane potential is converted into spike output at the initiation zone by active membrane conductances. We tested this directly by characterizing the mapping between the intracellular membrane potential close to the spike initiation zone and firing frequency. Because the LGMD is a large neuron, intracellular recordings can be made either in its dendritic fields or in the axon, close to the initiation zone. The site of impalement is determined by the position of the recording electrode, and can be verified after staining the cell and from the action potential waveform at the recording site (dendritic, Fig. 3a, b; axonal, Fig. 4a, b).

We first determined whether voltage-dependent sodium channels in (or close to) the axon participate in any other way than in converting membrane potential into instantaneous firing rate. To this end, we recorded the axonal, intracellular membrane potential (V_m) in response to approaching objects before and after local injection of the sodium channel blocker tetrodotoxin (TTX) proximal to the recording site (Fig. 4b, inset). Drug ejection was monitored by adding a dye to the ejected solution (Methods). This pharmacological manipulation abolished firing responses while leaving peak depolarizing synaptic potentials unaffected (Fig. 4b). In contrast, injections of TTX in the second optic chiasm onto presynaptic excitatory afferents to the LGMD abolished all membrane depolarizations (not shown). The control membrane potential ($\bar{V}_m$; in absence of TTX, Fig. 4b) was extracted by median filtering (Fig. 4a, inset, and 4b, orange lines; Methods). Comparison of TTX and control membrane potential traces revealed a considerable temporal shift (Fig. 4b), which we quantified by computing the average time difference between identical membrane potentials observed in the two conditions (average of the horizontal dotted lines' length in Fig. 4c). TTX-sensitive sodium conductances

advanced the response of the LGMD by 180 ± 110 ms (7 animals, 4 $l/|v|$ values per animal). Thus, TTX-sensitive conductances at or close to the spike initiation zone do not only serve to generate action potentials. They also normally accelerate the depolarization of the membrane potential during object approach.

Second, we measured the nonlinear transformation²⁸ between average postsynaptic potential at the axon (without pharmacological agents, after median filtering to remove spikes) and instantaneous firing rate. This nonlinear relationship is plotted in Fig. 4d for the experiment of Fig. 4a, b, averaged over different values of the looming stimulus parameter $l/|v|$. In this case, a third-order power law best described the transformation of membrane potential to firing rate (Fig. 4d). The same analysis was performed with 10 LGMD neurons, each from a different animal (Supplementary Table 1). In 7 out of 10 neurons (for example, Fig. 4d), a third-order power law best described the data. In one animal, no single-parameter model provided a good fit. In another, a sixth-order nonlinearity even closer to exponential than third order was required. In the last, exponential and third-power models could not be distinguished statistically.

Our experimental data suggest that two transformations occur between the LGMD membrane potential and firing rate during object approach. First is a temporal shift of the membrane potential time-course, resulting in a phase-advance of the spiking response (Fig. 4c). Because this shift represents a large fraction of the approach duration (>150 ms on average), it significantly amplifies motion-sensitive excitation. Together with the effective angular ranges of lateral and feedforward inhibition (Figs 2 and 3), this information places new constraints on the activation dynamics of presynaptic afferents to the LGMD during object approach. The second transformation is a nonlinear, expansive mapping of membrane potential into firing rate. Within the range of membrane potentials observed at the spike initiation zone, this transformation is best fitted by a third power that is remarkably close, but not identical, to an exponentiation (Fig. 4d). This suggests that either multiplication is only approximately implemented by a log-exp transform, or, alternatively, that part of the exponentiation occurs upstream of the spike initiation zone. Both transformations rely, at least in part, on TTX-sensitive channels.

In conclusion, we propose that the phenomenological fit between spatiotemporal features of approaching objects and the LGMD's firing rate—a product of two terms—represents the output of a dendritic and axonal implementation of multiplication by way of a log-exp transformation. We have suggested^{11,12} that multiplication could be carried out by the exponentiation of a sum of two postsynaptic currents or potentials: a positive excitatory one, representing the logarithm of angular velocity²⁶ and a negative one, representing angular size. We have shown here that the influence of angular velocity is excitatory and postsynaptic, and that the influence of angular size is inhibitory and postsynaptic. These results imply that multiplication occurs within LGMD itself, rather than presynaptically. Because the resulting axonal membrane potential maps onto output firing rate by way of an expansive nonlinearity close to an exponential, our experimental data are consistent with an approximate implementation by a log-exp transformation.

Multiplicative interactions between distinct sensory or extra-sensory variables have been reported in other preparations^{1–5}, although it has proved difficult either to identify the specific neurons responsible for these computations, or to map different sensory inputs onto distinct synaptic contributions to a single neuron. Thus the question of whether single neurons can implement multiplication remained largely open. Although several aspects of this computation remain to be characterized—in particular, the exact temporal dependence of feedforward excitation and inhibition on speed and size during object approach and their integration within the LGMD's dendritic tree—our results suggest

that the biophysical underpinning of such a nonlinear operation can be described mechanistically in a single neuron. □

Methods

Electrophysiology

Dissections and extracellular recordings from 45 descending contralateral movement detector (DCMD) neurons were as described in refs 12 and 20. We used extra-cellular recordings from the DCMD to monitor the spike activity of the LGMD because their spiking outputs match perfectly (Fig. 4a). Intracellular recordings from the LGMD ($n = 91$ animals) were obtained using glass microelectrodes filled with 2 M potassium acetate solution (30–80 MΩ). Typical intracellular recordings times were 40 min – 1 h. Stainings (Fig. 1a) were obtained by iontophoresis of Lucifer yellow (2% in aqueous solution). Local drug injections were performed using pulled and microforged glass micropipettes (final diameter, $\sim 1 \mu\text{m}$), backfilled with PCTX (5 mM) or TTX (1 μM) in aqueous solution and 0.5% (w/v) Fast green for visualization. Further technical details are available in Supplementary Information. Intracellular signals were amplified in bridge or DCC mode using an Axoclamp 2B (Axon Instruments) or SEL-10 (NPI) amplifier.

Visual stimulation

Visual stimulation procedures have been described previously^{12,20}. Briefly, black squares approaching perpendicular to the main body axis towards the eye (duration, 0.6–5.7 s) at various values of $l/|v|$ were presented in pseudo random order. Background flow stimuli consisted of alternating bright and grey (30% contrast) concentric rectangles moving outwards towards the screen boundaries at constant temporal frequency. For the slow-flow stimulus (Fig. 2a), the temporal frequency of light/dark transitions was set between 6 and 12 Hz. For the fast stimulus (Figs 2c, 3a, b, d), the temporal frequency was 50 Hz. Further details as well as procedures used to minimize habituation²⁹ of the responses are described in Supplementary information.

Data analysis

Data collection was performed using a PCI data acquisition card (UEI) and custom software²⁰. Sample rates were 10 kHz for extracellular, and 20 kHz for intracellular, recordings. Data analysis methods followed refs 12 and 20. Median filtering and fits of the LGMD/DCMD firing rate, $f(t)$, with various models are described in Supplementary Information. The relation between average membrane potential, $\langle V_m \rangle$, and firing rate, $\langle f \rangle$, (Fig. 4d) was obtained by averaging across trials membrane potential and instantaneous firing rate over 5-ms time windows and plotting one variable against the other. Linear ($\langle f \rangle = \alpha \langle V_m \rangle$), power law ($\langle f \rangle = \alpha \langle V_m \rangle^\mu$) and exponential ($\langle f \rangle = \exp(\alpha \langle V_m \rangle) - 1$) models depicted in Fig. 4d were fitted using maximum likelihood. Statistical tests followed ref. 30.

Received 1 August; accepted 20 September 2002; doi:10.1038/nature01190.

- Reichardt, W. Autokorrelations-Auswertung als Funktionsprinzip des Zentralnervensystems. *Z. Naturforsch.* **12**, 448–457 (1957).
- Barlow, H. B. & Levick, W. R. The mechanisms of directionally selective units in rabbit's retina. *J. Physiol. (Lond.)* **178**, 477–504 (1965).
- McAdams, C. J. & Maunsell, J. H. R. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J. Neurosci.* **19**, 431–441 (1999).
- Sun, H. & Frost, B. J. Computation of different optical variables of looming objects in pigeon nucleus retinundus neurons. *Nature Neurosci.* **1**, 296–303 (1998).
- Pena, J. L. & Konishi, M. Auditory spatial receptive fields created by multiplication. *Science* **292**, 249–252 (2001).
- Torre, V. & Poggio, T. A synaptic mechanism possibly underlying directional selectivity to motion. *Proc. R. Soc. Lond. B* **202**, 409–416 (1978).
- Andersen, R. A., Essick, G. K. & Siegel, R. M. Encoding of spatial location by posterior parietal neurons. *Science* **23**, 456–458 (1985).
- Yuste, R. & Tank, D. Dendritic integration in mammalian neurons, a century after Cajal. *Neuron* **16**, 701–716 (1996).
- Koch, C. *Biophysics of Computation: Information Processing in Single Neurons* (Oxford Univ. Press, Oxford, 1998).
- Chance, F. S., Abbott, L. F. & Reyes, A. D. Gain modulation from background synaptic input. *Neuron* **35**, 773–782 (2002).
- Hatsopoulos, N., Gabbiani, F. & Laurent, G. Elementary computation of object approach by a wide-field visual neuron. *Science* **270**, 1000–1003 (1995).
- Gabbiani, F., Krapp, H. G. & Laurent, G. Computation of object approach by a wide-field, motion-sensitive neuron. *J. Neurosci.* **19**, 1122–1141 (1999).
- Rind, F. C. & Bramwell, D. I. Neural network based on the input organization of an identified neuron signaling impending collision. *J. Neurophysiol.* **75**, 967–985 (1996).
- Rind, F. C. Intracellular characterization of neurons in the locust brain signaling impending collision. *J. Neurophysiol.* **75**, 986–995 (1996).
- Wagner, H. Flow field variables trigger landing in flies. *Nature* **297**, 147–148 (1982).
- Wickelmaier, M. & Strausfeld, N. J. Organization and significance of neurons that detect change of visual depth in the hawk moth *Manduca sexta*. *J. Comp. Neurol.* **424**, 356–376 (2000).
- Schlotterer, G. R. Response of the locust descending movement detector neuron to rapidly approaching and withdrawing visual stimuli. *Can. J. Zool.* **55**, 1372–1376 (1977).
- Rind, F. C. & Simmons, P. J. Orthopteran DCMD neuron: a reevaluation of responses to moving objects. I. Selective responses to approaching objects. *J. Neurophysiol.* **68**, 1654–1666 (1992).
- Robertson, R. M. & Johnson, A. G. Retinal image size triggers obstacle avoidance in flying locusts. *Naturwissenschaften* **80**, 176–178 (1993).
- Gabbiani, F., Mo, C. & Laurent, G. Invariance of angular threshold computation in a wide-field looming-sensitive neuron. *J. Neurosci.* **21**, 314–329 (2001).
- O'Shea, M. & Rowell, C. H. F. Protection from habituation by lateral inhibition. *Nature* **254**, 53–55 (1975).

- Rowell, C. H. F., O'Shea, M. & Williams, J. L. D. The neuronal basis of a sensory analyser, the acridid movement detector system. IV. The preference for small field stimuli. *J. Exp. Biol.* **68**, 157–185 (1977).
- Rind, F. C. & Simmons, P. J. Local circuits for the computation of object approach by an identified visual neuron in the locust. *J. Comp. Neurol.* **395**, 405–415 (1998).
- Rauh, J. J., Lummis, S. C. R. & Sattelle, D. B. Pharmacological and biochemical properties of insect GABA receptors. *Trends Pharmacol.* **11**, 325–329 (1990).
- Warzecha, A. K., Egelhaaf, M. & Borst, A. Neural circuit tuning fly visual interneurons to motion of small objects. I. Dissection of the circuit by pharmacological and photoinactivation techniques. *J. Neurophysiol.* **69**, 329–339 (1993).
- Zanker, J. M., Srinivasan, M. V. & Egelhaaf, M. Speed tuning in elementary motion detectors of the correlation type. *Biol. Cybern.* **80**, 109–116 (1999).
- Mead, C. *Analog VLSI and Neural Systems* (Addison-Wesley, Boston, 1989).
- Koch, C., Bernander, O. & Douglas, R. J. Do neurons have a voltage or a current threshold for action potential initiation? *J. Comput. Neurosci.* **2**, 63–82 (1995).
- Rowell, C. H. F. In *Experimental Analysis of Insect Behavior* (ed. Browne, L. B.) 87–99 (Springer, New York, 1974).
- Galant, A. R. *Non-linear Statistical Models* (Wiley & Sons, New York, 1987).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements We thank C. Mo for technical assistance with the experiments presented in Fig. 2, S. Potter for help with two-photon confocal microscopy, and J. Maunsell for comments. This work was supported by the Sloan Foundation, the National Institute of Mental Health (E.G., C.K.), the National Institute for Deafness and Communication Disorders (G.L.), the McKnight Foundation (G.L.) and the Center for Neuromorphic Engineering as part of the NSF Engineering Research Center programme. H.G.K. was supported by a travel grant of the Deutsche Forschungsgemeinschaft. F.G. is an Alfred P. Sloan research fellow.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to F.G. (e-mail: gabbiani@bcm.tmc.edu).

Essential role for TIRAP in activation of the signalling cascade shared by TLR2 and TLR4

Masahiro Yamamoto^{*,†}, Shintaro Sato^{*,†}, Hiroaki Hemmi^{*,†}, Hideki Sanjo^{*,†}, Satoshi Uematsu^{*,†}, Tsuneyasu Kaisho^{*,†,‡}, Katsuaki Hoshino^{*,†}, Osamu Takeuchi^{*,†}, Masaya Kobayashi^{*,†}, Takashi Fujita[§], Kiyoshi Takeda^{*,†} & Shizuo Akira^{*,†}

^{*} Department of Host Defense, Research Institute for Microbial Diseases, Osaka University, and [†] Solution Oriented Research for Science and Technology, Japan Science and Technology Corporation, 3-1 Yamada-oka, Suita, Osaka 565-0871, Japan

[‡] RIKEN Research Center for Allergy and Immunology, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

[§] Department of Tumor Cell Biology, The Tokyo Metropolitan Institute of Medical Science, 3-18-22 Honkomagome, Bunkyo-ku, Tokyo 113-8613, Japan

Signal transduction through Toll-like receptors (TLRs) originates from their intracellular Toll/interleukin-1 receptor (TIR) domain, which binds to MyD88, a common adaptor protein containing a TIR domain^{1–4}. Although cytokine production is completely abolished in MyD88-deficient mice, some responses to lipopolysaccharide (LPS), including the induction of interferon-inducible genes and the maturation of dendritic cells, are still observed^{5–7}. Another adaptor, TIRAP (also known as Mal), has been cloned as a molecule that specifically associates with TLR4 and thus may be responsible for the MyD88-independent response^{8,9}. Here we report that LPS-induced splenocyte proliferation and cytokine production are abolished in mice lacking TIRAP. As in MyD88-deficient mice, LPS activation of the nuclear factor NF- κ B and mitogen-activated protein kinases

is induced with delayed kinetics in TIRAP-deficient mice⁵. Expression of interferon-inducible genes and the maturation of dendritic cells is observed in these mice; they also show defective response to TLR2 ligands, but not to stimuli that activate TLR3, TLR7 or TLR9. In contrast to previous suggestions, our results show that TIRAP is not specific to TLR4 signalling and does not participate in the MyD88-independent pathway. Instead, TIRAP has a crucial role in the MyD88-dependent signalling pathway shared by TLR2 and TLR4.

To examine the physiological role of TIRAP, we generated TIRAP-deficient mice by gene targeting. A targeting vector was constructed in which the two exons encoding the TIR domain were replaced with the neomycin resistance (*neo*^R) gene (Fig. 1a). Homologous recombination in embryonic stem (ES) cells was confirmed by Southern blot analysis, and correctly targeted ES cells were used to generate mice with the mutated allele (Fig. 1b). Expression of TIRAP mRNA in the mutant mice was examined by northern blot analysis. When we used the amino-terminal fragment of TIRAP cDNA as a probe, we detected transcripts from the mutant mice with almost the same size as those of wild-type mice (Fig. 1c). Polymerase chain reaction with reverse transcription analysis (RT-PCR) showed, however, that the *neo*^R gene was inserted into the mouse *Tirap* gene in the mutant mice, leading to a stop codon in the N-terminal portion of TIRAP (Supplementary Fig. 1). Immunoblot analysis of splenocytes confirmed that disruption of the *Tirap* gene abolished the expression of TIRAP protein (Fig. 1d). TIRAP-deficient mice were healthy and had a normal composition of B and T lymphocytes, as determined by flow cytometry (data not shown).

Because TIRAP was identified as an adaptor that specifically associates with TLR4, we first analysed the response to LPS in TIRAP-deficient mice^{8,9}. Stimulation with LPS induced dose-dependent proliferation and augmented the expression of major

histocompatibility complex (MHC) class II molecules in wild-type splenocytes, whereas proliferation and augmentation of MHC class II expression was severely impaired in TIRAP-deficient splenocytes (Fig. 2a and b). Although TLR and the interleukin-1 (IL-1) receptor family share some signalling molecules, such as MyD88, both wild-type and TIRAP-deficient splenocytes proliferated to the same extent in response to IL-1 plus PHA, showing that TIRAP-deficient splenocytes are not defective in IL-1-induced proliferation (ref. 10 and Fig. 2c).

We analysed the LPS-induced production of inflammatory cytokines from peritoneal macrophages. Wild-type macrophages produced IL-6, tumour-necrosis factor- α (TNF- α) and IL-12 in response to LPS (Fig. 3a, b); however, LPS did not stimulate the production of IL-6, TNF- α or IL-12 in TIRAP-deficient macrophages. Wild-type mice showed an increased serum concentration of IL-6 after a high dose injection of LPS and died within 96 h (Fig. 3c and d). By contrast, TIRAP-deficient mice were completely resistant to LPS-induced shock and produced only a small amount of IL-6.

Wild-type and TIRAP-deficient macrophages produced similar amounts of IL-6, TNF- α and IL-12 in response to CpG DNA, a TLR9 ligand, and the synthetic compound R-848, a TLR7 ligand

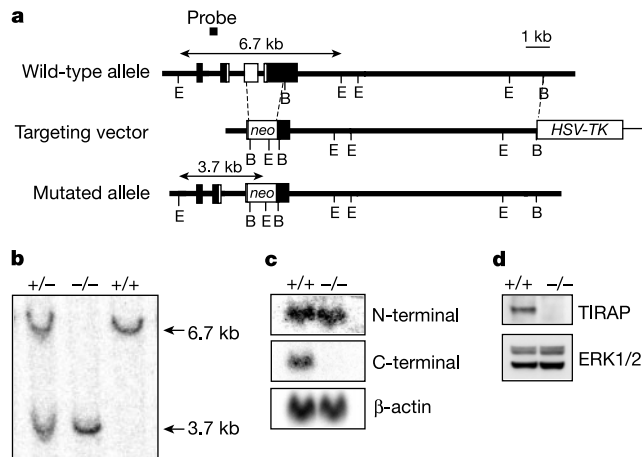


Figure 1 Targeted disruption of the murine *Tirap* gene. **a**, Structure of the mouse *Tirap* gene, the targeting vector and the predicted disrupted gene. Open boxes denote the coding exon. E, *EcoRI*; B, *Bam*HI. **b**, Southern blot analysis of offspring from the heterozygote intercrosses. Genomic DNA was extracted from mouse tails, digested with *EcoRI*, separated by electrophoresis and hybridized with the radiolabelled probe indicated in **a**. Southern blotting gave a single 6.7-kb band for wild-type (+/+), a 3.7-kb band for homozygous (–/–) and both bands for heterozygous (+/–) mice. **c**, Northern blot analysis of embryonic fibroblasts. Total RNA (10 μ g) extracted from embryonic fibroblasts was separated by electrophoresis, transferred to nylon membrane and hybridized using the TIRAP N-terminal or C-terminal fragment as a probe. The same membrane was rehybridized with a β -actin probe. **d**, Western blot analysis of splenocytes. The cell lysates were immunoprecipitated and immunoblotted with a polyclonal antibody against TIRAP. The same lysates were also blotted with antibodies against ERK1 and ERK2 to monitor protein expression.

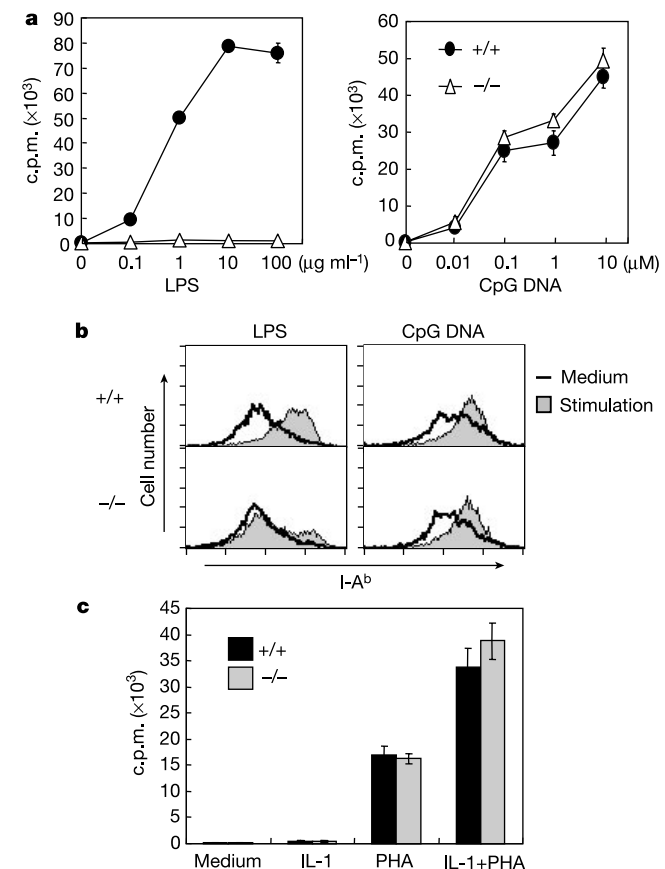


Figure 2 Impaired response to LPS in TIRAP-deficient splenocytes. **a**, Proliferation of splenocytes stimulated by LPS and CpG DNA. Wild-type (+/+) and TIRAP-deficient (–/–) splenocytes were cultured with the indicated concentration of LPS or CpG DNA for 48 h. [³H]thymidine (1 μ Ci) was pulsed for the last 12 h, and [³H]thymidine incorporation was measured by liquid scintillation counting. **b**, Splenic B cells were cultured with 1 μ g ml⁻¹ LPS or 0.1 μ M CpG DNA. After 48 h, cells were collected, stained with a biotinylated antibody against I-A^b followed by streptavidin-PE, and analysed by FACS. **c**, Splenocytes from wild-type and TIRAP-deficient mice were incubated with 10 ng ml⁻¹ IL-1 β in the presence of 10 ng ml⁻¹ PHA for 72 h. [³H]thymidine was pulsed for the last 12 h, and uptake of [³H] was measured.

(refs 11, 12, and Fig. 3a). Stimulation of macrophages with poly(I:C), a TLR3 ligand, induced the mRNA expression of interferon- β (IFN- β), as well as IL-6 and TNF- α , in both wild-type and TIRAP-deficient mice (ref. 13 and Fig. 3e).

We also analysed the production of inflammatory cytokines in response to the mycoplasma-derived lipopeptide MALP-2, and to *Staphylococcus aureus* peptidoglycan (PGN), which are both TLR2 ligands (refs 14, 15, and Fig. 3f). Wild-type macrophages showed dose-dependent production of IL-6, TNF- α and IL-12 in response to MALP-2 and PGN. In contrast, the MALP-2- or PGN-induced production of these cytokines was severely impaired in TIRAP-deficient mice. Together, these results show that TIRAP-deficient mice have an impaired response to TLR2 and TLR4 ligands, but a normal response to microbial components recognized by TLR3, TLR7 and TLR9.

Next, we analysed the activation of intracellular signalling cascades in response to LPS, MALP-2 and R-848. In wild-type macro-

phages, stimulation with LPS resulted in enhanced DNA-binding activity of NF- κ B (Fig. 4a, left). In TIRAP-deficient and MyD88-deficient macrophages, this NF- κ B DNA-binding activity was not induced at 10 min but was significantly induced at 20 min. Stimulation with LPS induced the activation of mitogen-activated protein (MAP) kinases, including extracellular-signal-regulated kinases (ERK) 1 and 2, p38 and Jun N-terminal kinase (JNK), in wild-type macrophages (Fig. 4b, left). In TIRAP-deficient macrophages, however, the LPS-induced activation of these kinases was detected with delayed kinetics. Thus, the LPS-induced activation of signalling molecules is compromised in TIRAP-deficient mice, whose phenotype is similar to that of MyD88-deficient mice⁵.

MALP-2 induced NF- κ B DNA binding activity in macrophages. However, NF- κ B activation was severely impaired in TIRAP-deficient mice (Fig. 4a, middle). In addition, the MALP-2-induced activation of ERK1/ ERK2, p38 and JNK was not observed in TIRAP-deficient macrophages (Fig. 4b, middle). By contrast, the

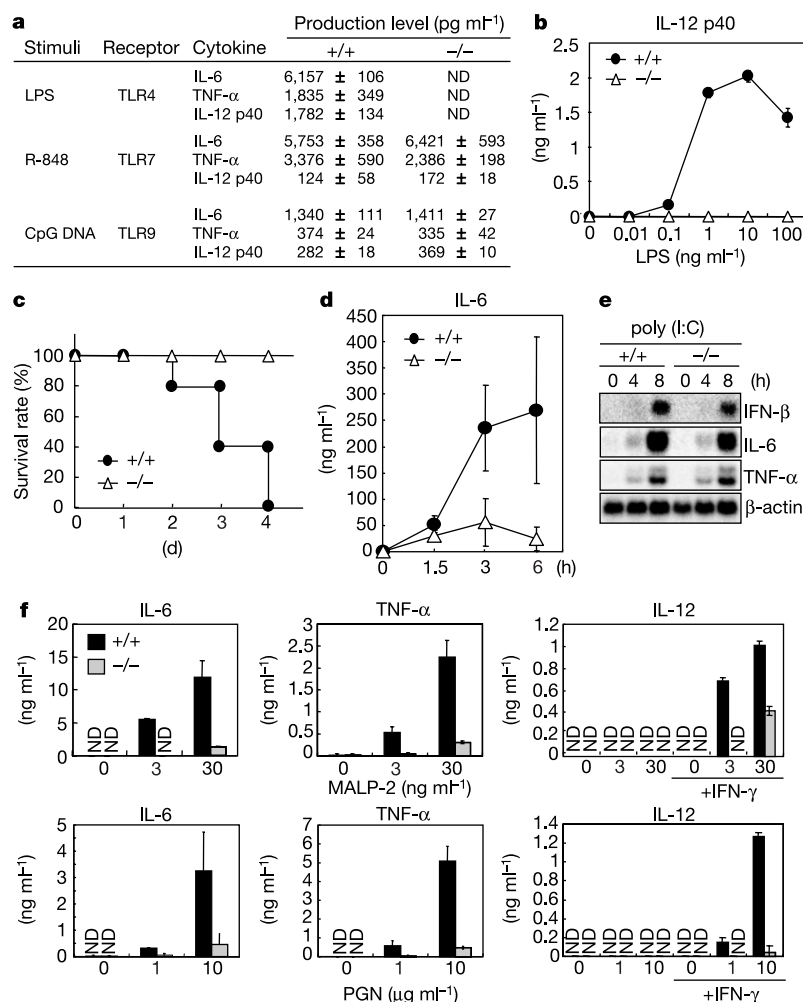


Figure 3 Cytokine production in peritoneal macrophages in response to TLR ligands. **a**, Peritoneal macrophages from wild-type (+/+) and TIRAP-deficient (-/-) mice were cultured with 1 ng ml⁻¹ LPS, 10 nM R-848 or 1 μ M CpG DNA in the presence (IL-12) or absence (IL-6 and TNF- α) of 30 ng ml⁻¹ IFN- γ for 24 h. Concentrations of IL-6, TNF- α and IL-12 in the culture supernatants were measured by enzyme-linked immunosorbent assay (ELISA). Data are the means $\pm$ s.d. of triplicates. ND, not detected. **b**, Peritoneal macrophages were stimulated with the indicated concentrations of LPS in the presence of 30 ng ml⁻¹ IFN- γ for 24 h, and the supernatants were analysed for IL-12 by ELISA. Data are the means $\pm$ s.d. of triplicates. **c**, Age-matched wild-type ($n = 5$) and TIRAP-deficient ($n = 5$) mice were injected intraperitoneally with 2.0 mg of

LPS (*Escherichia coli* O55:B5). Mortality was assessed daily for 4 d. **d**, TIRAP-deficient and wild-type mice were intraperitoneally injected with 1.5 mg of LPS (*E. coli* O55:B5). Sera were taken after the indicated period of time, and serum concentrations of IL-6 were determined by ELISA. Data are the means $\pm$ s.d. of samples from three mice. **e**, Peritoneal macrophages were stimulated with 50 μ g ml⁻¹ poly (I:C) for 4 and 8 h. Total RNA was extracted and analysed for expression of IFN- β , IL-6 and TNF- α mRNA by northern blotting. **f**, Peritoneal macrophages were stimulated with the indicated concentrations of MALP-2 or PGN for 24 h. Production of IL-6, TNF- α and IL-12 in the culture supernatants was measured by ELISA. Data are the means $\pm$ s.d. of triplicates. ND, not detected.

R-848-induced activation of NF- κ B and MAP kinases was not altered in TIRAP-deficient macrophages (Fig. 4a and b, right). We also analysed activity of IRAK-1 by *in vitro* kinase assay (Fig. 4c). In wild-type macrophages, stimulation with LPS, MALP-2 and R-848 induced autophosphorylation of IRAK-1. In TIRAP-deficient mice, however, IRAK-1 was not activated in response to either LPS or MALP-2, but was activated normally in response to R-848. Thus,

similar to MyD88, TIRAP acts at the level of IRAK-1 signalling; however, unlike activation mediated by MyD88, the TIRAP-mediated activation of IRAK-1 is specific to TLR2 and TLR4.

The TLR4 signal consists of the MyD88-dependent and -independent pathways^{1,5-7}. The MyD88-independent pathway contributes to the LPS-induced expression of IFN-inducible genes and the maturation of dendritic cells^{6,7}. *In vitro* studies have indicated that

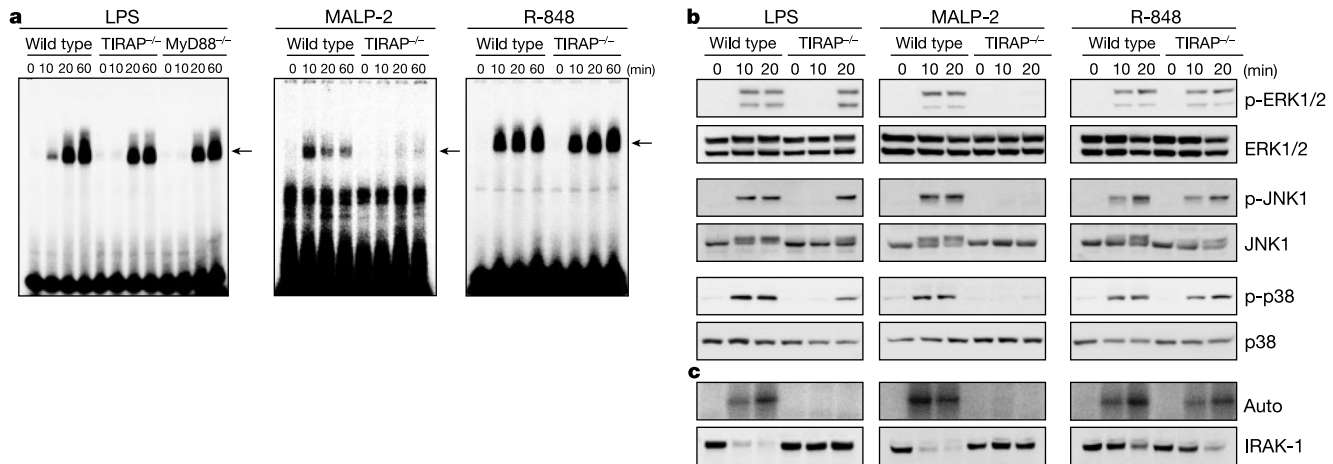


Figure 4 Activation of signalling cascades in response to TLR2, TLR4 and TLR7 ligands in peritoneal macrophages. **a**, Peritoneal macrophages were stimulated with 100 ng ml⁻¹ LPS (left), 3 ng ml⁻¹ MALP-2 (middle) or 10 nM R-848 (right) for the indicated durations. Nuclear extracts were prepared and NF- κ B DNA-binding activity was determined by electrophoretic mobility shift assay using a probe specific for NF- κ B. Arrows indicate the induced NF- κ B complex. **b**, Peritoneal macrophages were stimulated with LPS (left),

MALP-2 (middle) or R-848 (right) for 10 or 20 min. Whole-cell lysates were prepared and subjected to western blot analysis using antibodies specific for the indicated molecules. **c**, Lysates from macrophages stimulated with LPS (left), MALP-2 (middle) or R-848 (right) were immunoprecipitated with antibodies against IRAK-1. The kinase activity of IRAK-1 was measured by *in vitro* kinase assay (top). The same lysates were blotted with antibodies against IRAK-1 (bottom). Auto indicates auto-phosphorylation.

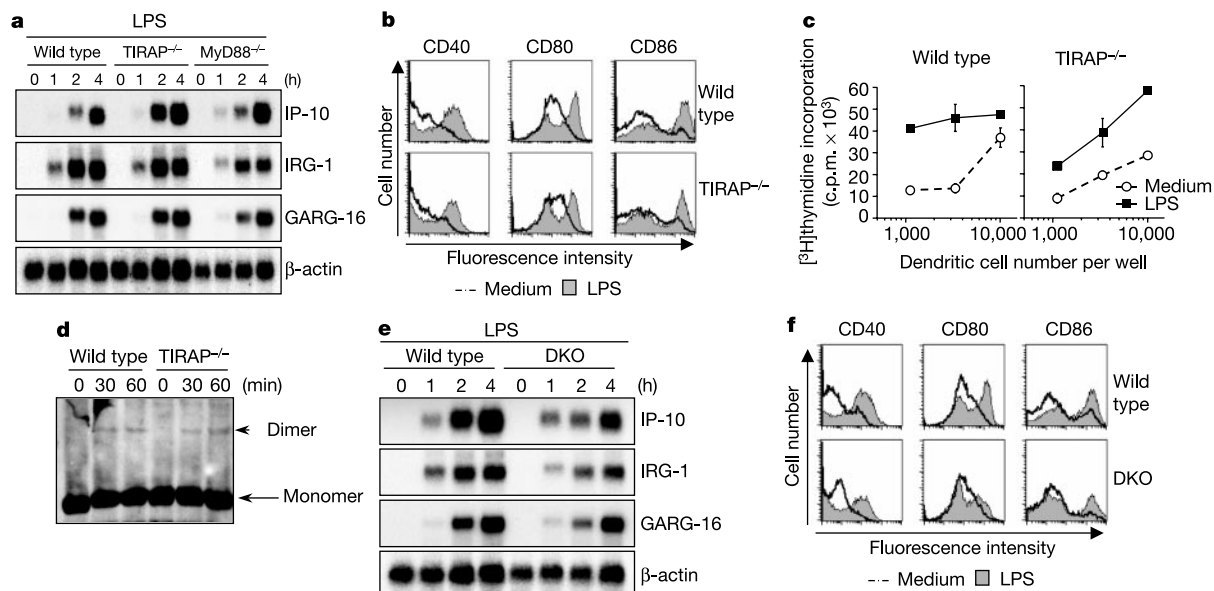


Figure 5 LPS-induced activation of the MyD88-independent signalling pathway in TIRAP-deficient mice. **a**, Peritoneal macrophages were stimulated with 1 ng ml⁻¹ LPS for the indicated durations. Total RNA (5 μ g) was extracted and subjected to northern blot analysis for expression of genes encoding IP-10, IRG-1 and GARG-16. The same membrane was re-hybridized with a β -actin probe. **b**, Bone-marrow-derived dendritic cells were stimulated with 100 ng ml⁻¹ LPS for 24 h and analysed for surface expression of the indicated molecules by flow cytometry. **c**, LPS-stimulated (squares) or unstimulated (open circles) bone-marrow-derived dendritic cells (H-2^b) were irradiated and incubated with allogeneic BALB/c CD4⁺ T cells (H-2^d) at the indicated concentrations. Proliferative

responses of T cells were evaluated by the [³H]thymidine incorporation of CD4⁺ T cells. **d**, Peritoneal macrophages were stimulated with 1 μ g ml⁻¹ LPS for 30 or 60 min. Cell lysates were prepared and resolved by native PAGE. Monomeric (arrow) and dimeric (arrowhead) forms of IRF-3 were detected by western blot. **e**, Peritoneal macrophages from wild-type and MyD88, TIRAP doubly deficient mice (DKO) were stimulated with LPS and analysed for the expression of IP-10, IP-10, IRG-1 and GARG-16 by northern blotting. **f**, Bone-marrow-derived dendritic cells from wild-type and MyD88, TIRAP doubly deficient mice (DKO) were cultured with 100 ng ml⁻¹ LPS for 24 h. The surface expression of the indicated molecules was analysed by flow cytometry.

TIRAP is involved in the LPS-induced activation of the MyD88-independent pathway^{8,9,16,17}. We therefore examined whether TIRAP deficiency affects the MyD88-independent signalling. Peritoneal macrophages from wild-type, TIRAP-deficient and MyD88-deficient mice were stimulated with LPS and analysed for the induction of IFN-inducible genes encoding IP-10, GARG-16 and IRG-1 by northern blot analysis (Fig. 5a). LPS induction of IP-10, GARG-16 and IRG-1 was comparable in wild-type, TIRAP-deficient and MyD88-deficient macrophages, indicating that the LPS induction of IFN-inducible genes was not impaired in TIRAP-deficient mice. LPS stimulation of wild-type dendritic cells (DCs) enhanced the expression of MHC class II and co-stimulatory molecules such as CD40, CD80 and CD86 (Fig. 5b). Similarly, TIRAP-deficient DCs showed enhancement of these surface molecules in response to LPS. In addition, LPS stimulation led to enhanced allo-stimulatory activity of DCs in both wild-type and TIRAP-deficient mice, showing that the LPS-induced maturation of DCs is retained in TIRAP-deficient mice, as it is in MyD88-deficient mice (Fig. 5c).

Because IRF-3 is reportedly activated in a MyD88-independent manner in response to LPS⁷, we analysed the LPS-induced activation of IRF-3. When wild-type macrophages were stimulated with LPS, IRF-3 was activated and formed homodimers, as assessed by native polyacrylamide gel electrophoresis (PAGE) analysis (Fig. 5d). The LPS-induced dimerization of IRF-3 was similarly observed in TIRAP-deficient macrophages. Thus, the LPS-induced activation of IRF-3 is not affected in the absence of TIRAP.

These results did not exclude the possibility that MyD88 might compensate for the TIRAP deficiency and vice versa during the LPS response. We therefore generated mice lacking both MyD88 and TIRAP. Genes encoding IP-10, IRG-1 and GARG-16 were induced normally in response to LPS in the MyD88, TIRAP doubly deficient macrophages (Fig. 5e), and LPS-induced expression of co-stimulatory molecules was observed normally in MyD88, TIRAP doubly deficient DCs (Fig. 5f). Taken together, these findings indicate that TIRAP may not be essential for the LPS-induced activation of the MyD88-independent pathway.

We have examined the functional role of TIRAP by gene targeting. Our results do not support a model in which TIRAP is responsible for the TLR4-specific MyD88-independent events in TLR4 signalling, including the induction of IFN-inducible genes and the maturation of dendritic cells^{8,9,16,17}. Instead, TIRAP has a role similar to that of MyD88 in TLR signalling. But the function of TIRAP seems to be confined to TLR4 and TLR2 signalling, which is distinct from the role of MyD88 as a common adaptor. Both MyD88 and TIRAP are essential to the common signalling pathway shared by TLR2 and TLR4, which leads to activation of IRAK-1 and NF- κ B, and finally to cytokine production. There might be additional adaptors involved in the signalling pathway through other TLR members, as well as in the LPS-stimulated MyD88-independent pathway. Identification of such molecules will be important to our understanding of the molecular mechanism of the specific biological effects elicited by individual TLRs. □

Methods

Generation of TIRAP-deficient mice

Genomic DNA containing the mouse *Tirap* gene was isolated from 129/SV mouse genomic library and characterized by restriction enzyme mapping and sequencing analysis. We constructed the targeting vector by replacing a 1.6-kb fragment encoding a part of TIR domain with a *neo^R* cassette, and inserted a herpes simplex virus thymidine kinase driven by the MCI promoter into the genomic fragment for negative selection (Fig. 1a). The targeting vector was transfected into embryonic stem cells (E14.1) G418 and gancyclovir doubly resistant colonies were selected and screened by PCR and Southern blotting. We micro-injected homologous recombinants into C57BL/6 female mice and intercrossed heterozygous F₁ progenies to obtain TIRAP-deficient mice. TIRAP-deficient mice and their wild-type littermates from these intercrosses were used for experiments.

Reagents

We purchased LPS from *Salmonella minnesota* Re 595 prepared by a phenol/chloroform/petroleum ether extraction procedure from Sigma, PGN from *S. aureus* from Fluka, and

poly(I:C) from Amersham. MALP-2 was synthesized and purified as described¹⁴. R-848 was a gift from the Pharmaceuticals and Biotechnology Laboratory of the Japan Energy Corporation. We prepared CpG oligodeoxynucleotides as described¹¹. The polyclonal antibody against TIRAP was obtained by immunizing rabbit with bacterially expressed His-tagged full-length murine TIRAP protein. Antibodies against phosphorylated ERK, JNK and p38 were purchased from Cell Signaling. Antibodies against ERK, JNK and p38 were from Santa Cruz.

Electrophoretic mobility shift assay

Peritoneal macrophages (1×10^6) were stimulated with 100 ng ml⁻¹ LPS, 3 ng ml⁻¹ MALP-2 and 10 nM R-848 for the indicated durations. We purified nuclear extracts from cells and incubated them with a probe specific for the NF- κ B DNA-binding site, separated them by electrophoresis and visualized them by autoradiography as described¹⁰.

Western blot analysis and *in vitro* kinase assay

Peritoneal macrophages were stimulated with 100 ng ml⁻¹ LPS, 3 ng ml⁻¹ MALP-2 or 100 nM R-848 for the indicated durations. The cells were then lysed in buffer containing 1.0% Nonidet-P40, 150 mM NaCl, 20 mM Tris-HCl (pH 7.5), 5 mM EDTA and protease inhibitor cocktail (Roche). We resolved the cell lysates by SDS-PAGE and transferred them onto a PVDF membrane (BioRad). The membrane was blotted with the indicated antibodies and visualized with an enhanced chemiluminescence system (NEN Life Science Product). We measured the IRAK-1 activity in cell lysates by *in vitro* kinase assay as described¹⁸.

Preparation and analysis of dendritic cells

Bone marrow cells from wild-type or TIRAP-deficient mice were cultured with 10 ng ml⁻¹ granulocyte-macrophage colony-stimulating factor for 6 d. Immature dendritic cells were collected and cultured in fresh medium in the absence or presence of 100 ng ml⁻¹ LPS for a further 24 h. We stained the dendritic cells first with biotinylated antibodies against CD40, CD80 or CD86 and then with phycoerythrin (PE)-conjugated streptavidin. We carried out flow cytometric analysis using a FACSCalibur with CELLQuest software (Becton Dickinson). The allo-stimulatory activity of dendritic cells was analysed as described⁶.

Native PAGE assay

Peritoneal macrophages were stimulated with 1 (μ g ml⁻¹) LPS for the indicated periods and then lysed. The cell lysates in native PAGE sample buffer (62.5 mM Tris-HCl (pH 6.8), 15% glycerol and 1% deoxycholate) were separated by native PAGE and then immunoblotted with antibodies against IRF-3 as described¹⁹.

Received 1 July; accepted 9 September 2002; doi:10.1038/nature01182.

1. Akira, S., Takeda, K. & Kaisho, T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nature Immunol.* **2**, 675–680 (2001).
2. Imer, J. L. & Hoffmann, J. A. Toll-like receptors in innate immunity. *Trends Cell Biol.* **11**, 304–311 (2001).
3. Medzhitov, R. Toll-like receptors and innate immunity. *Nature Rev. Immunol.* **1**, 135–145 (2002).
4. Janeway, C. A. Jr & Medzhitov, R. Innate immune recognition. *Annu. Rev. Immunol.* **20**, 197–216 (2002).
5. Kawai, T., Adachi, O., Ogawa, T., Takeda, K. & Akira, S. Unresponsiveness of MyD88-deficient mice to endotoxin. *Immunity* **11**, 115–122 (1999).
6. Kaisho, T., Takeuchi, O., Kawai, T., Hoshino, K. & Akira, S. Endotoxin-induced maturation of MyD88-deficient dendritic cells. *J. Immunol.* **166**, 5688–5694 (2001).
7. Kawai, T. *et al.* Lipopolysaccharide stimulates the MyD88-independent pathway and results in activation of IRF-3 and the expression of a subset of LPS-inducible genes. *J. Immunol.* **167**, 5887–5894 (2001).
8. Horng, T., Barton, G. M. & Medzhitov, R. TIRAP: an adapter molecule in the Toll signaling pathway. *Nature Immunol.* **2**, 835–841 (2001).
9. Fitzgerald, K. A. *et al.* Mal (MyD88-adaptor-like) is required for Toll-like receptor-4 signal transduction. *Nature* **413**, 78–83 (2001).
10. Adachi, O. *et al.* Targeted disruption of the MyD88 gene results in loss of IL-1- and IL-18-mediated function. *Immunity* **9**, 143–150 (1998).
11. Hemmi, H. *et al.* A toll-like receptor recognizes bacterial DNA. *Nature* **408**, 740–745 (2000).
12. Hemmi, H. *et al.* Small anti-viral compounds activate immune cells via the TLR7-MyD88-dependent signaling pathway. *Nature Immunol.* **3**, 196–200 (2002).
13. Alexopoulou, L., Holt, A. C., Medzhitov, R. & Flavell, R. A. Recognition of double-stranded RNA and activation of NF- κ B by Toll-like receptor 3. *Nature* **413**, 732–738 (2001).
14. Takeuchi, O. *et al.* Cutting edge: preferentially the R-stereoisomer of the mycoplasma lipopeptide macrophage-activating lipopeptide-2 activates immune cells through a toll-like receptor 2- and MyD88-dependent signaling pathway. *J. Immunol.* **164**, 554–557 (2000).
15. Takeuchi, O. *et al.* Differential roles of TLR2 and TLR4 in recognition of Gram-negative and Gram-positive bacterial cell wall components. *Immunity* **11**, 443–451 (1999).
16. Toshchakov, V. *et al.* TLR4, but not TLR2, mediates IFN- β -induced STAT1 α/β -dependent gene expression in macrophages. *Nature Immunol.* **3**, 392–398 (2002).
17. O'Neill, L. A. Toll-like receptor signal transduction and the tailoring of innate immunity: a role for Mal? *Trends Immunol.* **23**, 296–300 (2002).
18. Sato, S. *et al.* A variety of microbial components induce tolerance to lipopolysaccharide by differentially affecting MyD88-dependent and -independent pathways. *Int. Immunol.* **14**, 783–791 (2002).
19. Iwamura, T. *et al.* Induction of IRF-3/-7 kinase and NF- κ B in response to double-stranded RNA and virus infection: common and unique pathways. *Genes Cells* **6**, 375–388 (2001).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements We thank P. F. Muhlradt and H. Tomizawa for MALP-2 and R-848, respectively; E. Horita for secretarial assistance; and N. Okita and N. Iwami for technical assistance. This work was supported by grants from Special Coordination Funds, the Ministry of Education, Culture, Sports, Science and Technology, and Research Fellowships from the Japan Society for the Promotion of Science for Young Scientists.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.A. (e-mail: sakira@biken.osaka-u.ac.jp).

The adaptor molecule TIRAP provides signalling specificity for Toll-like receptors

Tiffany Horng, Gregory M. Barton, Richard A. Flavell & Ruslan Medzhitov

Howard Hughes Medical Institute, Section of Immunobiology, Yale University School of Medicine, New Haven, Connecticut 06520, USA

Mammalian Toll-like receptors (TLRs) function as sensors of infection and induce the activation of innate and adaptive immune responses^{1–3}. Upon recognizing conserved pathogen-associated molecular products, TLRs activate host defence responses through their intracellular signalling domain, the Toll/interleukin-1 receptor (TIR) domain, and the downstream adaptor protein MyD88 (refs 1–3). Although members of the TLR and the interleukin-1 (IL-1) receptor families all signal through MyD88, the signalling pathways induced by individual receptors differ. TIRAP, an adaptor protein in the TLR signalling pathway, has been identified and shown to function downstream of TLR4 (refs 4, 5). Here we report the generation of mice deficient in the *Tirap* gene. TIRAP-deficient mice respond normally to the TLR5, TLR7 and TLR9 ligands, as well as to IL-1 and IL-18, but have defects in cytokine production and in activation of the nuclear factor NF- κ B and mitogen-activated protein kinases in response to lipopolysaccharide, a ligand for TLR4. In addition, TIRAP-deficient mice are also impaired in their responses to ligands for TLR2, TLR1 and TLR6. Thus, TIRAP is differentially involved in signalling by members of the TLR family and may account for specificity in the downstream signalling of individual TLRs.

Studies have established the TLR family as the essential recognition and signalling component of mammalian host defence^{1–3}. Ten mammalian TLRs have been described so far, and microbial ligands for many of them have been identified. TLR ligands represent molecular products derived from all of the main classes of pathogens and include the TLR4 ligand lipopolysaccharide (LPS)⁶, the TLR3 ligand double-stranded RNA⁷, the TLR2 ligands peptidoglycan (PGN)⁸ and bacterial lipoproteins^{9–11}, the TLR5 ligand flagellin¹², and the TLR9 ligand unmethylated CpG DNA¹³. Synthetic ligands for TLR7, imiquimod and resiquimod (R-848), have also been described¹. Understanding the functional differences between individual TLRs, especially in terms of signalling pathways and cellular responses induced by various microbial ligands, is a subject of great interest^{3,14}.

The adaptor protein TIRAP may be responsible for differential signalling by individual TLRs^{4,5}. A dominant-negative form of TIRAP inhibits the activation of nuclear factor NF- κ B induced by TLR4, but not by TLR9 or the interleukin 1 receptor (IL-1R)⁴. Consistent with this, TIRAP interacts with TLR4 but not with TLR9 in glutathione S-transferase pull-down assays. On the basis of these

results, we considered that TIRAP might be responsible for the cellular responses that can be observed in MyD88-deficient cells stimulated by LPS but not CpG. Specifically, although LPS-stimulated MyD88-deficient dendritic cells (DCs) can no longer make cytokines, they still retain the ability to upregulate the co-stimulatory molecules CD80 and CD86 (ref. 15). In addition, NF- κ B and mitogen-activated protein (MAP) kinases are still activated, albeit with delayed kinetics, in MyD88-deficient macrophages^{16,17}. Similar to TLR9, TLR2 is thought to lack signalling through the MyD88-independent pathway, because macrophage-activating lipopeptide 2Kd (MALP-2), a ligand for the TLR2–TLR6 receptor pair, does not activate NF- κ B in MyD88-deficient macrophages¹⁶.

To characterize the role of TIRAP in innate immunity, we generated TIRAP-deficient mice by homologous recombination. Embryonic stem (ES) cells were electroporated with the targeting vector, in which a 640-base-pair (bp) exon encoding most of the coding sequence of the *Tirap* gene was replaced with the neomycin resistance (*neo*^R) cassette (Fig. 1a). Two independently derived, correctly targeted ES cell clones (Fig. 1b) were then microinjected into C57BL/6 blastocysts. Chimaeric pups were bred to female C57BL/6 mice, and transmission of the mutated allele was monitored by Southern blot analysis of genomic DNA from the offspring (data not shown). Heterozygous mice were then interbred to produce offspring homozygous for the *Tirap* deletion (Fig. 1c). TIRAP-deficient mice were born at the expected mendelian ratio and showed no obvious gross abnormalities. Experiments were done with both knockout lines and similar results were obtained. The absence of TIRAP mRNA was verified by polymerase chain reaction with reverse transcription (RT–PCR) analysis of RNA isolated from the knockout mice (Fig. 1d).

We first looked at the effects of the *Tirap* deletion in B cells. Splenocyte proliferation was measured in response to the TLR4 ligand LPS, the TLR7 ligand R-848, the TLR9 ligand CpG and tripalmitoyl cysteinyl (Pam3Cys) lipopeptide (bacterial lipopeptide; BLP), which signals through the TLR1–TLR2 heterodimer¹⁸.

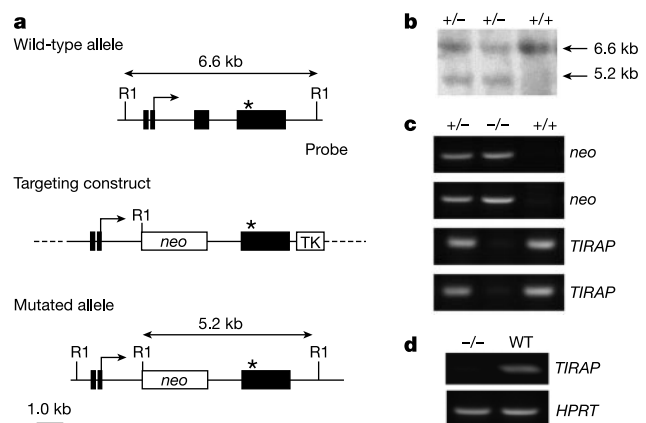


Figure 1 Generation of TIRAP-deficient mice. **a**, The mouse *Tirap* gene, the targeting vector and the mutated allele. Filled boxes denote exons, and the arrow and the asterisk mark the AUG translation start site and the stop codon, respectively. 'R1' marks the *Eco*RI sites used for Southern blot analysis. TK, thymidine kinase gene. **b**, Southern blot analysis of *Eco*RI-digested genomic DNA from the two injected ES cell clones (left and middle) and a wild-type ES cell clone (right). **c**, PCR of genomic tail DNA from homozygous knockout (–/–), heterozygous (+/–) and homozygous wild-type (+/+) mice using two combinations each of primers that detect the *neo*^R gene (top two panels) and the deleted exon (bottom two panels). **d**, RT–PCR of cDNA from TIRAP-deficient and wild-type (WT) littermates indicate the absence of TIRAP mRNA in TIRAP-deficient cells (top). Amplification of *HPRT* cDNA was included as a control (bottom).

TIRAP-deficient splenocytes showed a defect in proliferation in response to LPS but not to CpG or R-848 (Fig. 2), consistent with the *in vitro* studies that implicated TIRAP in TLR4 but not TLR9 signalling. These data show that TIRAP has a role in B-cell proliferation in response to LPS, although the defect is less pronounced than in MyD88-deficient splenocytes. Notably, TIRAP-deficient splenocytes were also impaired in their responsiveness to BLP (Fig. 2), indicating that TIRAP also has a role in TLR2 signalling.

We assessed the ability of TIRAP-deficient splenocytes to upregulate co-stimulatory molecules in response to LPS. Unlike MyD88-deficient B cells, which failed to upregulate CD86, TIRAP-deficient B cells were not impaired in their ability to induce CD86 upon LPS stimulation (see Supplementary Information). TIRAP is therefore involved in regulating the proliferation but not the induction of co-stimulatory molecules in B cells, whereas MyD88 is required for both.

We next analysed the phenotype of TIRAP-deficient bone-marrow-derived DCs. Production of IL-6, IL-12 and/or tumour-necrosis factor- α (TNF- α) by DCs stimulated with LPS, CpG, R-848 or BLP was measured by enzyme-linked immunoabsorbent assay (ELISA). TIRAP-deficient DCs produced wild-type amounts of these cytokines in response to CpG and R-848, but this production was significantly reduced in response to LPS and BLP (Fig. 3a and Supplementary Information). This defect in the TIRAP-deficient DCs is comparable to that of MyD88-deficient DCs, which produce barely detectable amounts of cytokines (ref. 15 and Supplementary Information). A similar defect in cytokine production was observed in TIRAP-deficient macrophages stimulated with LPS and BLP, but not with CpG (see Supplementary Information).

We assessed the ability of TIRAP-deficient DCs to mature in response to TLR ligands and found that they upregulated CD86 and CD40 in response to LPS and CpG (Fig. 3b). This is in contrast to MyD88-deficient DCs, which can mature only in response to LPS and not CpG. Therefore, both MyD88 and TIRAP regulate the pathway that leads to cytokine production, but the roles of MyD88

and TIRAP in inducing co-stimulatory molecules on LPS-stimulated DCs is not yet clear; it is possible that neither TIRAP-deficient nor MyD88-deficient DCs show a defect in maturation because the two adaptors can compensate for each other.

We next tested whether TIRAP is involved in TLR5 signalling. The TLR5 ligand flagellin induces an increase in the serum concentration of IL-6 after its injection into mice¹². We found that there was no difference in the serum concentrations of IL-6 and IL-12 in wild-type and TIRAP-deficient mice after they were injected with flagellin. By contrast, flagellin does not induce IL-6 production in MyD88-deficient mice¹². Therefore, TIRAP does not seem to be involved in TLR5-induced cytokine production (see Supplementary Information).

Our previous *in vitro* studies indicated that TIRAP, unlike MyD88, may not be involved in signalling downstream of the IL-1R family^{4,5}. To confirm this, we analysed the ability of TIRAP-deficient thymocytes to proliferate in response to IL-1. We found no reduction in the proliferative capacity of TIRAP-deficient cells, whereas MyD88-deficient cells failed to proliferate (Fig. 4a). In addition, TIRAP-deficient splenocytes produced normal amounts of interferon- γ (IFN- γ) in response to IL-18 as assessed by ELISA (Fig. 4b). Therefore, TIRAP may not be involved in signalling downstream of IL-1R and IL-18R and, unlike other previously described components of the TLR signalling pathway, is restricted in its function to a subset of the TLR family.

In the absence of MyD88, NF- κ B and the MAP kinases Jun amino-terminal kinase (JNK) and p38 are still activated in response to LPS, albeit with delayed kinetics¹⁷. We therefore assessed the ability of TIRAP-deficient macrophages to activate NF- κ B in response to TLR2, TLR4 and TLR9 ligands. When stimulated with LPS, TIRAP-deficient macrophages activated NF- κ B with delayed kinetics, very similar to the activation observed in MyD88-deficient cells (Fig. 5a). The same delay was observed also with TIRAP-deficient DCs (data not shown). Notably, the

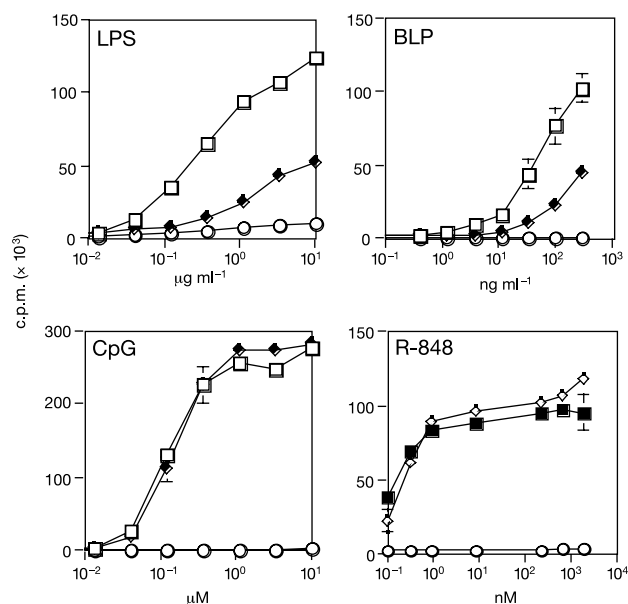


Figure 2 Impaired proliferation of TIRAP-deficient splenocytes to LPS and BLP but not CpG and R-848. Splenocytes from wild-type (squares), TIRAP-deficient (diamonds) and MyD88-deficient (circles) mice were stimulated in triplicate with titrations of LPS, CpG, R-848 or BLP for 36 h, and then pulsed with [^3H] thymidine for 12 h before cell collection. Data are representative of three independent experiments.

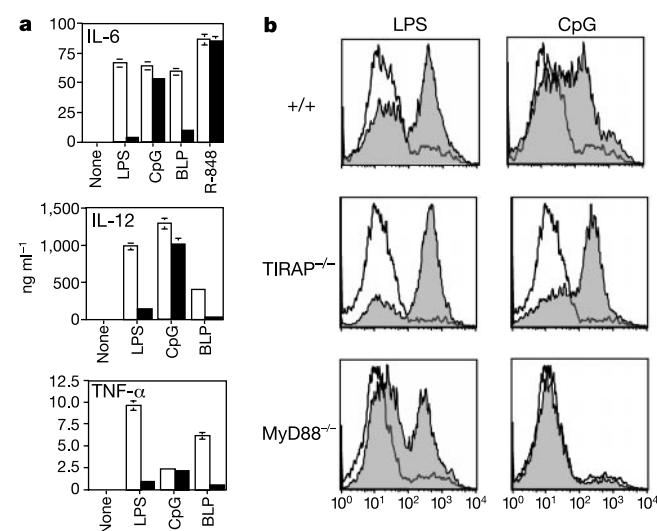


Figure 3 TIRAP-deficient DCs are significantly impaired in LPS- and BLP-induced cytokine production. **a**, Day 5 DCs from TIRAP-deficient mice (filled bars) or wild-type littermates (open bars) were left unstimulated or stimulated with LPS (10 ng ml^{-1}), CpG ($5 \mu\text{M}$), BLP (100 ng ml^{-1}) or R-848 (100 nM). Supernatants were collected for IL-6, IL-12 and/or TNF- α analysis by ELISA 36 h later. **b**, Day 6 DCs from wild-type (+/+), TIRAP-deficient and MyD88-deficient mice left untreated (open) or stimulated overnight (filled) with LPS (10 ng ml^{-1}) or CpG ($5 \mu\text{M}$) were stained with fluorescein isothiocyanate (FITC)-conjugated antibodies against CD11c and phycoerythrin (PE)-conjugated antibodies against CD86, and then analysed by FACS. The plots shown are gated on CD11c $^{+}$ cells. Data are representative of three independent experiments.

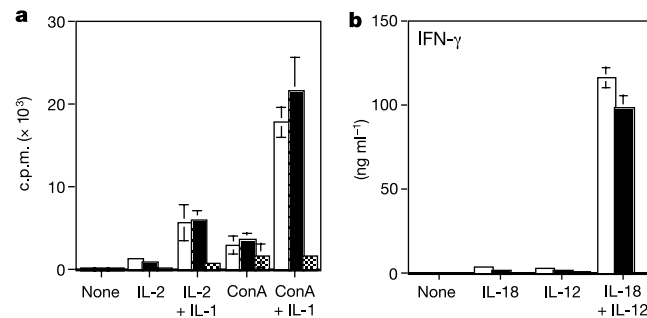


Figure 4 TIRAP is not involved in IL-1 and IL-18 signalling. **a**, Thymocytes from TIRAP-deficient (black bars), wild-type (white bars) and MyD88-deficient (grey bars) mice were stimulated with IL-2 (2 ng ml $^{-1}$) or ConA (0.625 μ g ml $^{-1}$) either alone or in combination with IL-1 (10 ng ml $^{-1}$) for 72 h. Cells were pulsed with [3 H]thymidine for 12 h before collection. **b**, Splenocytes from TIRAP-deficient (black bars), wild-type (white bars) and

MyD88-deficient (grey bars) mice were stimulated with either IL-18 (10 ng ml $^{-1}$) or IL-12 (10 ng ml $^{-1}$) alone, or with both IL-18 (10 ng ml $^{-1}$) and IL-12 (1 ng ml $^{-1}$) for 48 h, and then the supernatants were collected for analysis of IFN- γ by ELISA. Data are representative of three independent experiments.

activation of NF- κ B in TIRAP-deficient macrophages was both delayed and significantly reduced in response to BLP (Fig. 5a), which is consistent with the defects that we observed in BLP-induced B-cell proliferation and cytokine production by DCs. MyD88-deficient macrophages did not activate NF- κ B in response to BLP, as expected from previous results with MALP-2, another TLR2 ligand¹⁶. Notably, activation of NF- κ B in response to CpG was

not affected in the absence of TIRAP, which is consistent with the unimpaired CpG-induced cytokine production and B-cell proliferation.

We next examined the activation of MAP kinases in TIRAP-deficient macrophages in response to LPS, BLP and CpG. We found that activation of JNK, p38 and ERK in TIRAP-deficient macrophages occurred with delayed kinetics in response to LPS and BLP

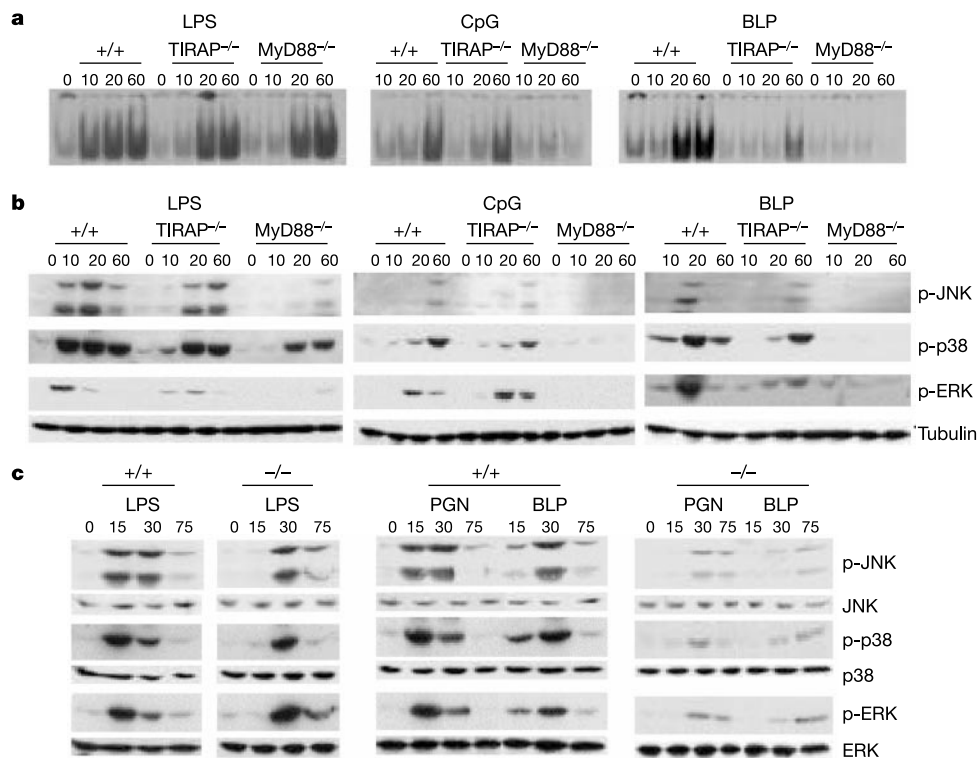


Figure 5 TIRAP-deficient macrophages activate NF- κ B and MAP kinases with delayed kinetics in response to LPS and BLP but not CpG. **a**, Wild-type (+/+), TIRAP-deficient and MyD88-deficient bone-marrow-derived macrophages were stimulated with LPS (100 ng ml $^{-1}$), CpG (10 μ M) or BLP (250 ng ml $^{-1}$) for the indicated durations. Activation of NF- κ B was assessed by electrophoretic mobility shift assay. Data are representative of three independent experiments. **b**, LPS (100 ng ml $^{-1}$), CpG (10 μ M) or BLP (250 ng ml $^{-1}$) was added to wild-type, TIRAP-deficient and MyD88-deficient bone-marrow-derived

macrophages for the indicated durations. Activation of JNK, p38 and ERK was analysed using phospho-specific antibodies. Analysis of α -tubulin was carried out as a loading control. **c**, Bone-marrow-derived macrophages from wild-type, TIRAP-deficient and MyD88-deficient mice were stimulated as indicated with LPS (100 ng ml $^{-1}$), PGN (100 μ g ml $^{-1}$) or BLP (250 ng ml $^{-1}$). Activation of JNK, p38 and ERK, as indicated, was assessed as above. Immunoblotting with antibodies to total JNK, P38 and ERK was done as a loading control.

(Fig. 5b). CpG-induced activation of these MAP kinases was not affected, consistent with our results for NF- κ B activation. By contrast, the activation of the MAP kinases in MyD88-deficient macrophages was delayed in response to LPS (ref. 17 and Fig. 5b), but completely absent in cells stimulated with BLP or CpG (Fig. 5b). Because the defect in TLR2 signalling had not been anticipated, we examined the ability of TIRAP-deficient macrophages to respond to peptidoglycan (PGN), which unlike BLP signals through the TLR2–TLR6 heterodimer¹⁹. As with BLP, PGN induced TIRAP-deficient macrophages to activate MAP kinases with delayed kinetics (Fig. 5c), suggesting that the attenuated responses of TIRAP-deficient cells to BLP may be generalized to other TLR2 ligands.

Our results show that different TLRs have varying requirements for MyD88 and TIRAP, and underscore the different roles of the two adaptors in TLR signalling. MyD88 is required for TLR9 signalling, because all CpG-induced responses (cytokine production, DC maturation, NF- κ B and MAP kinase activation, and B-cell proliferation) are abolished in the absence of MyD88. MyD88 is also required for signalling by TLR2, which cannot activate NF- κ B and MAP kinases in MyD88-deficient cells. By contrast, TLR4 can induce DC maturation and the activation of NF- κ B and MAP kinases in the absence of MyD88. It is therefore surprising that TLR2 signalling, like TLR4 signalling, is impaired in the absence of TIRAP, whereas TLR7, TLR9 and IL-1R family signalling are completely independent of TIRAP, in particular because TLR9 signalling is more similar to signalling by TLR2 and TLR4 in many respects (such as DC maturation and cytokine production), than it is to signalling by IL-1R and IL-18R. Our results indicate that the ability to induce MyD88-independent signalling is not only related to the involvement of TIRAP, although mice deficient in both TIRAP and MyD88 will need to be analysed to show this formally. That is, if both TIRAP and MyD88 can each induce DC maturation independently, then a defect in DC maturation will only be apparent in TIRAP, MyD88 doubly deficient mice. Alternatively, it is possible that a distinct protein, not TIRAP, is responsible for activating the MyD88-independent signalling pathway downstream of TLR4.

Regardless of these possibilities, our data clearly indicate that the MyD88-dependent signalling pathway is sufficient for the downstream signalling of some TLRs (for example, TLR7 and TLR9) and IL-1Rs, but not that of other TLRs (for example, TLR2–TLR1, TLR2–TLR6 and TLR4), which also require TIRAP. In summary, we have provided genetic evidence that TIRAP is an important component of innate host defense. TIRAP is differentially involved in signalling by members of the TLR family, and as such may be one of the signalling factors that accounts for the specificity in the downstream signalling of individual TLRs. □

Methods

Generation of TIRAP-deficient mice

The *Tirap* gene was isolated from ICR Swiss mice by chromosome walking using the GenomeWalker kit (Clontech). We constructed a targeting vector from a 3.7-kilobase (kb) *Clal*–*EcoRV* fragment containing the 5' end of the *Tirap* gene and a 3.1-kb *Sall*–*XhoI* fragment containing the 3' end of the *Tirap* gene. All but 70 bp at the 5' end and 20 bp at the 3' end of the *Tirap* gene are encoded on one exon, which was replaced in the targeting construct by the *neo*^R cassette. The targeting vector also included the thymidine kinase gene, which was used for negative selection of clones with random integration of the targeting vector. The targeting construct was electroporated into W9.5 ES cells, and neomycin-resistant, thymidine-kinase-insensitive clones were screened by Southern blot analysis (Fig. 1b) for evidence of homologous recombination.

We injected two independently derived ES cell clones into C57BL/6 blastocysts. Chimaeric male offspring were mated to female C57BL/6 mice, and transmission of the mutated allele in the heterozygous pups was determined by screening genomic DNA by Southern blot analysis (data not shown). Southern blot analysis was done according to standard procedures on *EcoRI*-digested genomic DNA using a 560-bp 3' fragment as the external probe. Heterozygotes were interbred to produce the homozygous TIRAP-deficient mice and wild-type littermates used for these studies. We screened the offspring of heterozygous breeding by PCR analysis of genomic DNA.

Cell culture

To obtain bone-marrow-derived macrophages, we cultured bone marrow cells for 7 d in complete RPMI media supplemented with supernatant containing macrophage colony-stimulating factor (M-CSF). Cells were fed every 2 d, and collected and replated on day 7 for use the next day. To obtain bone-marrow-derived DCs, we cultured bone marrow cells for 5 d in RPMI supplemented with granulocyte-macrophage CSF. On day 5, cells were either counted, replated into 48-well plates (5×10^5 cells per well in 0.5 ml) and stimulated with TLR ligands for analysis of cytokine production by ELISA, or stimulated directly in the well for staining and FACS analysis the next day.

Reagents

Tripalmitoyl cysteinyl (Pam3Cys) lipopeptide (BLP) was purchased from Bachem, LPS from Sigma, and PGN from Fluka. CpG DNA was synthesized by the Keck Facility at Yale University. All antibodies for western blotting were from NEB Cell Signaling. Antibodies against CD86, CD11c and CD40 for FACS were from Pharmingen. We purchased rIL-1 from R&D Systems, rIL-18 from MBL, rIL-12 and rTNF- α from BD Biosciences, and rIFN- γ and rIL-6 from Pharmingen. All antibodies for cytokine capture and detection for ELISAs were from Pharmingen.

Electrophoretic mobility shift assay

Bone-marrow-derived macrophages or DCs were collected and pooled, and then stimulated with the respective TLR ligands for the indicated durations. We prepared nuclear extracts and carried out gel-shift analysis as described²⁰.

Western blotting

Bone-marrow-derived macrophages plated in 60-mm tissue-culture dishes at 2.5×10^6 cells per dish were stimulated as indicated. We prepared cell lysates as described⁴ and quantified them by protein assay (Pierce). Lysates were analysed by SDS–PAGE followed by immunoblotting as described⁴.

Proliferation assays

For B-cell proliferation assays, erythrocyte-depleted splenocytes were incubated with LPS, CpG or BLP in 96-well plates (2×10^5 cells per well). After 36 h, proliferation was measured by [³H]thymidine incorporation for an additional 12–16 h. For thymocyte proliferation assays, we stimulated 2×10^5 thymocytes with IL-2 or concanavalin A (Con A), either individually or together with IL-1. Proliferation was measured as described above.

Received 9 July; accepted 6 September 2002; doi:10.1038/nature01180.

1. Akira, S., Takeda, K. & Kaisho, T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nature Immunol.* **2**, 675–680 (2001).
2. Adorem, A. & Ulevitch, R. J. Toll-like receptors in the induction of the innate immune response. *Nature* **406**, 782–787 (2000).
3. Medzhitov, R. Toll-like receptors and innate immunity. *Nature Rev. Immunol.* **1**, 135–145 (2001).
4. Horng, T., Barton, G. & Medzhitov, R. TIRAP, an adapter molecule in the Toll signaling pathway. *Nature Immunol.* **2**, 835–841 (2001).
5. Fitzgerald, K. A. *et al.* Mal (MyD88-adaptor-like) is required for Toll-like receptor-4 signal transduction. *Nature* **413**, 78–83 (2001).
6. Poltorak, A. *et al.* Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in *Tlr4* gene. *Science* **282**, 2085–2088 (1998).
7. Alexopoulos, L., Holt, A. C., Medzhitov, R. & Flavell, R. A. Recognition of double-stranded RNA and activation of NF- κ B by Toll-like receptor 3. *Nature* **413**, 732–738 (2001).
8. Takeuchi, O. *et al.* Differential roles of TLR2 and TLR4 in recognition of Gram-negative and Gram-positive bacterial cell wall components. *Immunity* **11**, 443–451 (1999).
9. Aliprantis, A. O. *et al.* Cell activation and apoptosis by bacterial lipoproteins through toll-like receptor-2. *Science* **285**, 736–739 (1999).
10. Brightbill, H. D. *et al.* Host defense mechanisms triggered by microbial lipoproteins through toll-like receptors. *Science* **285**, 732–736 (1999).
11. Underhill, D. M., Ozinsky, A., Smith, K. D. & Adorem, A. Toll-like receptor-2 mediates mycobacteria-induced proinflammatory signaling in macrophages. *Proc. Natl Acad. Sci. USA* **96**, 14459–14463 (1999).
12. Hayashi, F. *et al.* The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* **410**, 1099–1103 (2001).
13. Hemmi, H. *et al.* Toll-like receptor recognizes bacterial DNA. *Nature* **408**, 740–745 (2000).
14. Underhill, D. M. & Ozinsky, A. Toll-like receptors: key mediators of microbe detection. *Curr. Opin. Immunol.* **14**, 103–110 (2002).
15. Kaisho, T., Takeuchi, O., Kawai, T., Hoshino, K. & Akira, S. Endotoxin-induced maturation of myd88-deficient dendritic cells. *J. Immunol.* **166**, 5688–5694 (2001).
16. Kawai, T. *et al.* Lipopolysaccharide stimulates the MyD88-independent pathway and results in activation of IFN-regulatory factor 3 and the expression of a subset of lipopolysaccharide-inducible genes. *J. Immunol.* **167**, 5887–5894 (2001).
17. Kawai, T., Adachi, O., Ogawa, T., Takeda, K. & Akira, S. Unresponsiveness of MyD88-deficient mice to endotoxin. *Immunity* **11**, 115–122 (1999).
18. Takeuchi, O. *et al.* Cutting edge: role of toll-like receptor 1 in mediating immune response to microbial lipoproteins. *J. Immunol.* **169**, 10–14 (2002).
19. Ozinsky, A. *et al.* The repertoire for pattern recognition of pathogens by the innate immune system is defined by cooperation between toll-like receptors. *Proc. Natl Acad. Sci. USA* **97**, 13766–13771 (2000).
20. Kopp, E. & Ghosh, S. Inhibition of NF- κ B by sodium salicylate and aspirin. *Science* **265**, 956–959 (1994).

Supplementary Information accompanies the paper on Nature's website
(<http://www.nature.com>).

Acknowledgements We thank S. Akira for the MyD88-deficient mice; L. Kopp and C. Pasare for discussion and reading the manuscript; C. Annicelli for maintaining the mouse colony; L. Evangelista for technical assistance; L. Alexopoulou and K. Kobayashi for technical advice; and the Yale Cancer Center Transgenic Mouse and Gene Targeting Resource for blastocyst injections. This work was supported by the NIH, the Searle Foundation (R.M.) and the Howard Hughes Medical Institute (R.M., T.H., G.B. and R.F.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.M. (e-mail: ruslan.medzhitov@yale.edu).

A central role for JNK in obesity and insulin resistance

Jiro Hirosumi^{1,2}, Gürol Tuncman^{1,2}, Lufen Chang², Cem Z. Görgün¹, K. Teoman Uysal¹, Kazuhisa Maeda¹, Michael Karin² & Gökhan S. Hotamisligil¹

¹ Division of Biological Sciences and Department of Nutrition, Harvard School of Public Health, 665 Huntington Avenue, Boston, Massachusetts 02115, USA

² Laboratory of Gene Regulation and Signal Transduction, Department of Pharmacology, University of California San Diego, School of Medicine, 9500 Gilman Drive, La Jolla, California 92093, USA

† These authors contributed equally to this work

Obesity is closely associated with insulin resistance and establishes the leading risk factor for type 2 diabetes mellitus, yet the molecular mechanisms of this association are poorly understood¹. The c-Jun amino-terminal kinases (JNKs) can interfere with insulin action in cultured cells² and are activated by inflammatory cytokines and free fatty acids, molecules that have been implicated in the development of type 2 diabetes^{3,4}. Here we show that JNK activity is abnormally elevated in obesity. Furthermore, an absence of JNK1 results in decreased adiposity, significantly improved insulin sensitivity and enhanced insulin receptor signalling capacity in two different models of mouse obesity. Thus, JNK is a crucial mediator of obesity and insulin resistance and a potential target for therapeutics.

Obesity and type 2 diabetes are the most prevalent and serious metabolic diseases; they affect more than 50% of adults in the USA⁵. These conditions are associated with a chronic inflammatory response characterized by abnormal cytokine production, increased acute-phase reactants and other stress-induced molecules³. Many of these alterations seem to be initiated and to reside within adipose tissue, an unusual site for inflammation³. Elevated production of tumour necrosis factor (TNF)- α by adipose tissue decreases sensitivity to insulin and has been detected in several experimental obesity models and obese humans^{6,7}. Free fatty acids (FFAs) are also implicated in the aetiology of obesity-induced insulin resistance, although the molecular pathways involved in their action remain unclear⁴. Because both TNF- α and FFAs are potent JNK activators^{8–11}, we asked whether obesity is associated with alterations in stress-activated and inflammatory responses through this signalling pathway and whether JNKs are causally linked to aberrant metabolic control in this state.

We examined JNK activity in liver, muscle and adipose tissues of various models of obesity compared with lean controls to determine whether obesity activates the JNK pathway. In both dietary and genetic (*ob/ob*) models of obesity, there was a significant increase in total JNK activity in all tissues tested (Fig. 1a). In these tissues there was no apparent difference in the expression of either JNK1 or JNK2 polypeptides, suggesting that the activity of one or both of these

kinases is increased in response to obesity.

To test the functional significance of this alteration in the pathogenesis of obesity, insulin resistance and type 2 diabetes, we induced obesity in mice lacking either JNK1 (*Jnk1*^{−/−}) or JNK2 (*Jnk2*^{−/−}). *Jnk1*^{−/−} or *Jnk2*^{−/−} mice and their control littermates (*Jnk1*^{+/+} or *Jnk1*^{+/-} and *Jnk2*^{+/+} or *Jnk2*^{+/-}) were placed on a high-fat (50% of total calories derived from fat) and high-caloric diet (5286 kcal kg^{−1}; Bioserve, Frenchtown, NJ, USA) along with a control group of each genotype on a standard diet. On the high-fat diet, both controls and *Jnk2*^{−/−} mice developed marked obesity in comparison with mice kept on standard diet (Fig. 1b and c). Weight gain in *Jnk2*^{+/+}, *Jnk2*^{+/-} and *Jnk2*^{−/−} animals was indistinguishable on either standard or high-fat diet. However, weight gain on both standard and high-fat diets was significantly decreased for the *Jnk1*^{−/−} group (Fig. 1d and e). Although animals with one targeted allele of *Jnk1* (*Jnk1*^{+/-}) had a body weight intermediate between that of wild-type and *Jnk1*^{−/−} mice maintained on either diet, these differences did not reach statistical significance (Fig. 1e).

We assessed whether these differences in weight gain were related to alterations in adiposity. Sections of adipose tissue from *Jnk1*^{−/−} mice exhibited decreased adipocyte size relative to wild-type controls (Fig. 2a). This was not observed in *Jnk2*^{−/−} adipose tissue. The fat pad weights of *Jnk1*^{+/+}, *Jnk1*^{+/-} and *Jnk1*^{−/−} mice were similar in the lean group at both subcutaneous and epididymal fat depots (data not shown). However, in the obese group, the average weight

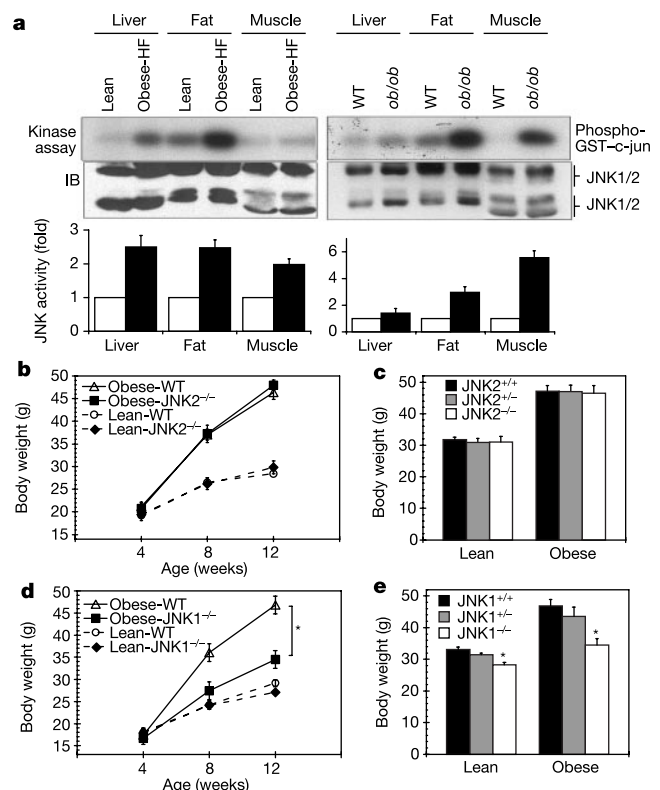


Figure 1 Total JNK activity and development of diet-induced obesity. **a**, Total JNK activity and protein concentrations in liver, muscle and adipose tissues of lean and obese mice. Obese HF, dietary model (high-fat); *ob/ob*, genetic model for leptin deficiency. In JNK immunoblots, JNK1 and JNK2 have relative molecular masses of 56,000–54,000 and 46,000–43,000, respectively. Lower panels show means $\pm$ s.e.m. of the quantified and normalized activities. **b–e**, Development of diet-induced obesity in *Jnk2*^{−/−} (**b, c**) and *Jnk1*^{−/−} (**d, e**) mice. All mice were male, 16 weeks old and on C57Bl/6J background. In **b** and **d**, $n = 10$ in each group. **c, e**, Means $\pm$ s.e.m. of body weights of male mice. Asterisk, statistical significance ($P < 0.05$) in a two-tailed Student *t*-test comparing *Jnk1*^{−/−} or *Jnk2*^{−/−} mice with controls.

of the subcutaneous fat depot was decreased by 33% in *Jnk1*^{-/-} mice compared with wild-type controls (Fig. 2b). Surprisingly, the weight of the epididymal fat pad was higher in the obese *Jnk1*^{-/-} group than in the wild-type controls, indicating a redistribution of adipose depots (Fig. 2b). No difference in fat pad weight was evident between *Jnk2*^{-/-} and wild-type mice in either condition (data not shown). To investigate systemic alterations in adiposity, we next examined total body composition. These studies demonstrated significantly decreased total body adiposity in *Jnk1*^{-/-} mice compared with controls (Fig. 2c). In contrast, the body composition of the *Jnk2*^{-/-} group was indistinguishable from wild-type controls (data not shown).

To address alternative causes for decreased body weight in *Jnk1*^{-/-} mice, we compared lipid metabolism, food intake, intestinal lipid absorption and core body temperature of *Jnk1*^{-/-} and *Jnk1*^{+/+} mice. No significant differences were observed in plasma triglyceride, cholesterol and FFA concentrations between genotypes (data not shown). Examination of faecal lipid content also did not reveal

any differences, thus excluding changes in intestinal lipid absorption (Fig. 2d). There was a small and statistically insignificant decrease in daily food intake (0.46 g d⁻¹) and increase (0.32 °C) in core body temperature in obese *Jnk1*^{-/-} mice compared with wild-type mice (Supplementary Fig. 1a and b). Although we cannot rule out the possibility that these small changes might contribute to decreased weight gain, the results strongly suggest that the deficiency in JNK1 is associated primarily with decreased adipocyte size, decreased adiposity and adipose redistribution in the context of dietary obesity.

Adipose tissue can have a substantial impact on systemic glucose homeostasis through production of bioactive molecules. We examined serum concentrations of adipocyte-derived secreted proteins with postulated roles in obesity and insulin action¹²⁻¹⁵. ACRP30 (30-kDa adipocyte complement-related protein)/adiponectin concentrations in the obese *Jnk1*^{-/-} mice were significantly higher than in *Jnk1*^{+/+} controls (Fig. 2e). In contrast, the concentrations of resistin were lower in *Jnk1*^{-/-} mice than in *Jnk1*^{+/+} animals (Fig. 2f). Because recent studies have indicated a role for adiponectin as a mediator of fatty-acid oxidation and hepatic insulin

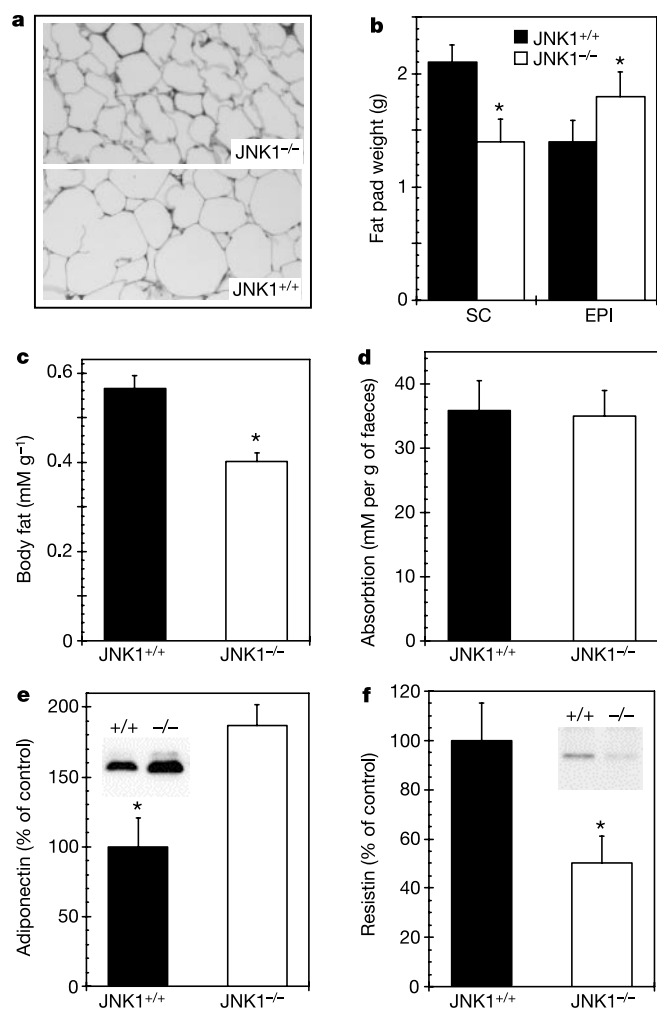


Figure 2 Adipose tissue morphology and adiposity in *Jnk1*^{-/-} mice and wild-type controls. **a, b**, Histological sections of epididymal fat pads (original magnification ×50) (**a**) and subcutaneous (SC) and epididymal (EPI) fat pad weights (**b**) of 16-week-old male *Jnk1*^{-/-} and *Jnk1*^{+/+} mice ($n = 3$ in **a**, $n = 9$ in **b**). **c-f**, Total body composition (**c**), faecal lipid content (**d**), serum adiponectin concentration (**e**) and resistin concentration (**f**) in *Jnk1*^{-/-} and *Jnk1*^{+/+} mice. Representative immunoblots are shown in insets. Total carcass lipid analysis was performed²⁵ to determine fat mass of individual mice ($n = 6$ in each group). Asterisk, statistical significance ($P < 0.05$) in a two-tailed Student *t*-test comparing *Jnk1*^{+/+} and *Jnk1*^{-/-} mice.

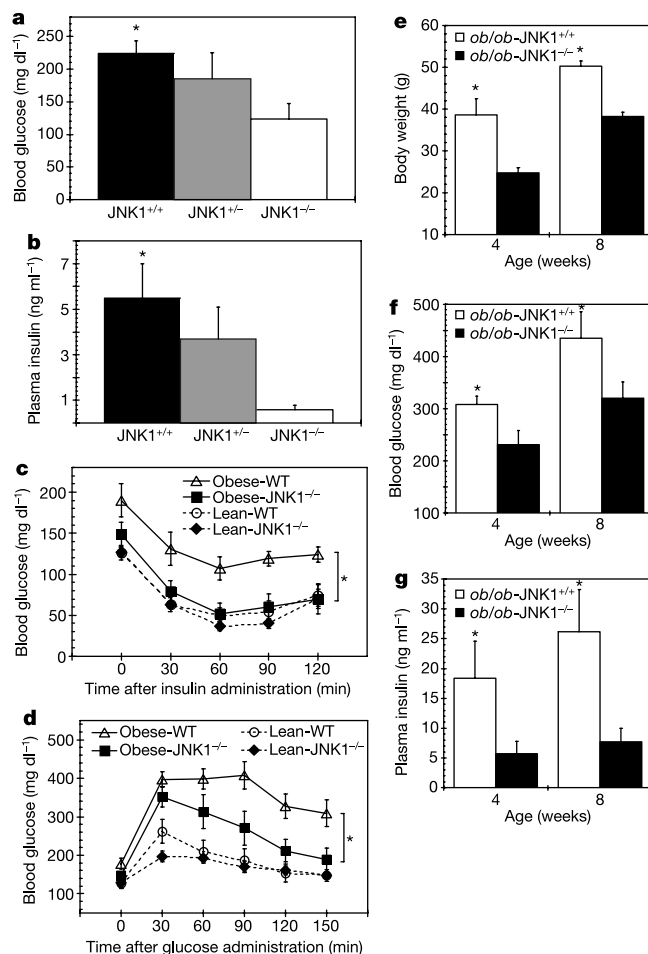


Figure 3 Metabolic effects of JNK1 deficiency. **a-d**, Examination of glucose homeostasis by fasting plasma glucose (**a**) and insulin (**b**) concentrations and insulin (**c**) and glucose (**d**) tolerance tests in lean and obese *Jnk1*^{-/-} and control male mice at 16 weeks of age. Investigation of the dynamics of the responses to the tolerance tests were done by analysis of variance repeated-measures analysis (Statview 4.01, Abacus Concepts, Berkeley, CA, USA). **e-g**, Body weight (**e**) and blood glucose (**f**) and plasma insulin (**g**) in *ob/ob-Jnk1*^{+/+} and *ob/ob-Jnk1*^{-/-} mice. Body weight and blood measurements for *ob/ob* mice were performed at 4 and 8 weeks of age and after a 6-h daytime food withdrawal. Asterisk, statistical significance ($P < 0.001$ in **c** and **d**; $P < 0.05$ in **e-g**). WT, wild type.

sensitivity^{12,13} and resistin is postulated to have a role in insulin resistance¹⁴, these alterations could affect systemic insulin sensitivity.

To test this possibility, we investigated glucose homeostasis in *Jnk1*^{-/-} and *Jnk2*^{-/-} mice and in wild-type controls. Obese *Jnk1*^{+/+} mice developed mild hyperglycaemia compared with lean wild-type controls (224 ± 20 versus 126 ± 11 mg dl⁻¹ (mean $\pm$ standard error of the mean; s.e.m.), $P < 0.001$). In contrast, obese *Jnk1*^{-/-} mice had significantly lower blood glucose concentrations than obese *Jnk1*^{+/+} mice (Fig. 3a). At 12 weeks of age, the blood glucose concentration in obese *Jnk1*^{-/-} mice was indistinguishable from that of lean *Jnk1*^{+/+} or *Jnk1*^{-/-} animals (148 ± 15 versus 126 ± 11 and 127 ± 8 mg dl⁻¹, $P = 0.8$). Obese wild-type mice also developed significant hyperinsulinaemia compared with those on the standard diet (5.5 ± 1.5 versus 0.69 ± 0.1 ng ml⁻¹, $P < 0.001$). Blood insulin in obese *Jnk1*^{-/-} mice was significantly lower than

in obese *Jnk1*^{+/+} controls (Fig. 3b) and was indistinguishable from either *Jnk1*^{+/+} or *Jnk1*^{-/-} lean mice (0.63 ± 0.18 versus 0.69 ± 0.16 and 0.57 ± 0.13 ng ml⁻¹, $P = 0.8$). Blood glucose and insulin concentrations in *Jnk1*^{+/+} mice were intermediate between those of *Jnk1*^{+/+} and *Jnk1*^{-/-} animals, but these differences were statistically insignificant (Fig. 3a and b). Obese *Jnk2*^{-/-} mice developed a similar degree of hyperglycaemia and hyperinsulinaemia to that in obese wild-type animals. Blood glucose and insulin concentrations were indistinguishable between the *Jnk2*^{-/-}, *Jnk2*^{+/-} and *Jnk2*^{+/+} groups (Supplementary Fig. 2a and b). The increase in blood glucose and insulin in animals on the high-fat diet is a strong indicator of obesity-induced insulin resistance and progression to type 2 diabetes. These results therefore indicate that the JNK1- but not JNK2-deficient animals are protected from the development of obesity-induced insulin resistance.

To investigate this point further, we performed intraperitoneal insulin and glucose tolerance tests (IITT and IGTT, respectively). The hypoglycaemic response to insulin was lower in obese *Jnk1*^{+/+} mice throughout the experiment than in obese *Jnk1*^{-/-} animals (Fig. 3c). Again, the glucose disposal curves of obese *Jnk1*^{-/-} mice were indistinguishable from those of lean animals. Integration of the area under the glucose disposal curves illustrated an overall difference of 40% between *Jnk1*^{+/+} and *Jnk1*^{-/-} mice (Supplementary Fig. 3a and b). IGTT also revealed a higher degree of hyperglycaemia in obese *Jnk1*^{+/+} animals throughout the experiment than in obese *Jnk1*^{-/-} mice (Fig. 3d). However, in this test the responses recorded in obese *Jnk1*^{-/-} mice did not reach those of lean controls, especially in the early phases, indicating residual insulin resistance (Fig. 3d). Interestingly, increased responsiveness in IGTT was even evident in lean *Jnk1*^{-/-} mice at the early phase of the experiment. In contrast, obese *Jnk2*^{-/-} animals exhibited marked insulin resistance in both IITT and IGTT (Supplementary Fig. 4a and b). The response curves of obese *Jnk2*^{-/-} mice were essentially identical to those of obese wild-type animals. In summary, both tests confirm that the ablation of *Jnk1* substantially decreases the development of insulin resistance associated with dietary obesity.

We next generated genetically obese mice (*ob/ob*) with targeted mutations in *Jnk1* to test the action of JNK1 in a different and more severe model of obesity. As expected, *ob/ob* mice developed early-onset and severe obesity (Fig. 3e). However, the extent of weight gain was lower in *ob/ob-Jnk1*^{-/-} mice than in *ob/ob-Jnk1*^{+/+} animals. Furthermore, at both 4 and 8 weeks of age the blood glucose concentrations were lower in *ob/ob-Jnk1*^{-/-} mice than in *ob/ob-Jnk1*^{+/+} animals (Fig. 3f). The *ob/ob-Jnk1*^{+/+} animals displayed a severe and progressive hyperinsulinaemia (18.4 ± 6.2 and 26.4 ± 7.1 ng ml⁻¹ at 4 and 8 weeks of age, respectively). In contrast, plasma insulin concentrations of *ob/ob-Jnk1*^{-/-} were lower (5.7 ± 2.1 and 7.7 ± 2.3 ng ml⁻¹ at 4 and 8 weeks of age, respectively) (Fig. 3g). Thus, JNK1 deficiency can provide partial resistance against obesity, hyperglycaemia and hyperinsulinaemia even in the most severe form of the disease associated with leptin deficiency in *ob/ob* mice.

In many but not all functions mediated by JNK, redundancy and molecular compensation were observed^{8,16-20}. To seek a mechanistic explanation for the involvement of JNK1 isoforms in obesity-related insulin resistance, we examined total JNK activity in liver, muscle and adipose tissues of obese *Jnk1*^{-/-} and *Jnk2*^{-/-} mice and in obese wild-type controls. These experiments demonstrated that JNK1 deficiency significantly decreases the obesity-induced increase in total JNK activity at all sites examined (Fig. 4a). No such decrease was observed in *Jnk2*^{-/-} mice (data not shown). Similar observations were also made after treatment of wild-type, *Jnk1*^{-/-} and *Jnk2*^{-/-} mice with lipopolysaccharide and using wild-type, *Jnk1*^{-/-} and *Jnk2*^{-/-} mouse embryo fibroblasts (data not shown). Thus, the JNK1 isoforms account for most, if not all, of the increased total JNK activity in the target tissues relevant for obesity-induced insulin resistance. This might be related to the specific activity of JNK1 or

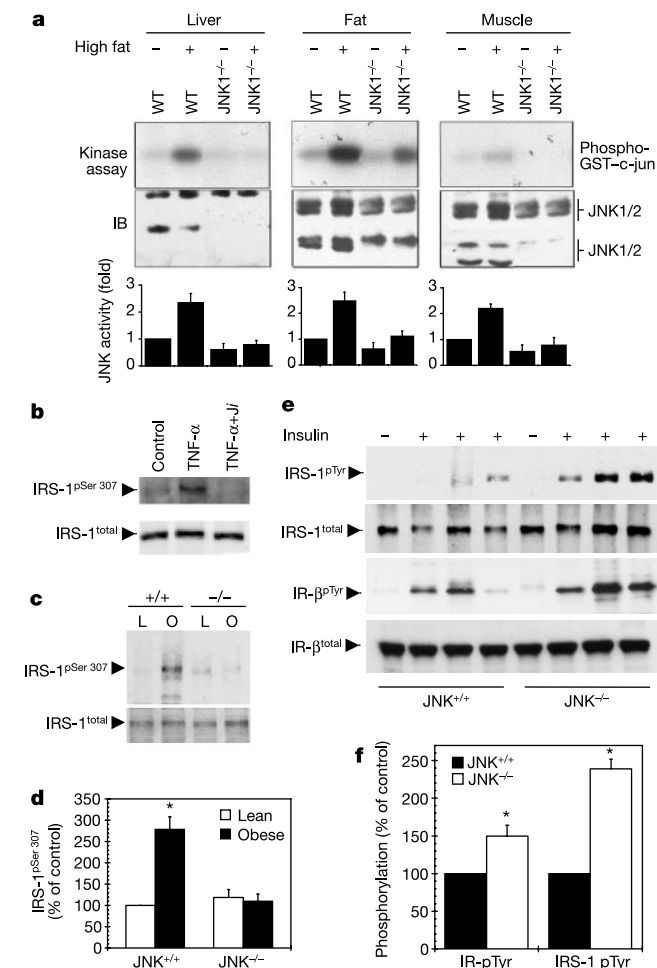


Figure 4 JNK activity and insulin signalling in JNK1-deficient mice. **a**, JNK activity and protein concentrations in liver, muscle and adipose tissues of lean and obese *Jnk1*^{+/+} (WT) and *Jnk1*^{-/-} mice. In JNK immunoblots, JNK1 and JNK2 have relative molecular masses of 56,000–54,000 and 46,000–43,000, respectively. **b**, Phosphorylation of IRS-1 at Ser 307 in liver cells treated for 1 h with 10 ng ml⁻¹ TNF- α in the absence (control) or presence of a specific JNK inhibitor SP600125 (*Ji*) (ref. 26) at 2.5 μ M. **c–f**, Phosphorylation of IRS-1 at Ser 307 (**c**, **d**) and insulin receptor (IR) signalling (**e**, **f**) in obese *Jnk1*^{-/-} and *Jnk1*^{+/+} mice. Total and Ser 307-phosphorylated IRS-1 concentrations were determined in liver tissues from lean (L) and obese (O) mice. Representative immunoblots of insulin-stimulated tyrosine phosphorylation (pTyr) of IR and IRS-1 in liver tissues of *Jnk1*^{-/-} and *Jnk1*^{+/+} mice are shown in **e**. Each lane represents an individual mouse. Graphs in **a**, **d** and **f** show means $\pm$ s.e.m. of the immunoblots. WT, wild type.

the relative abundances of the two isoforms in target tissues.

We next examined potential molecular mechanisms that might underlie the protection from insulin resistance conferred by the loss of *Jnk1*. Inhibitory serine phosphorylation of insulin receptor substrate (IRS)-1 was previously shown to be responsible for both TNF- α -induced and FFA-induced insulin resistance^{21,22}. Direct involvement of JNK in insulin signalling was also suggested, on the basis of experiments *in vitro*, to be exerted through phosphorylation of IRS-1 at Ser 307 (ref. 2). We examined whether this mechanism is involved in the action of JNK1 on IRS-1 and obesity-induced insulin resistance *in vivo*. We found increased IRS-1 phosphorylation at Ser 307 in a cellular model of insulin resistance in liver cells treated with TNF- α (Fig. 4b). TNF- α -induced Ser 307 phosphorylation was completely prevented by a JNK inhibitor (Fig. 4b). This TNF- α treatment regimen resulted in a significant decrease in insulin-stimulated tyrosine phosphorylation of IRS-1 (data not shown). We also examined IRS-1 Ser 307 phosphorylation in liver tissue of lean and obese *Jnk1*^{+/+} and *Jnk1*^{-/-} mice. The extent of IRS-1 Ser 307 phosphorylation was markedly increased in obese wild-type mice relative to the lean controls (Fig. 4c and d). Most importantly, no such increase could be detected in obese *Jnk1*^{-/-} mice, demonstrating that Ser 307 of IRS-1 is a target for JNK action *in vivo* (Fig. 4c and d). Finally, we found that insulin-induced IRS-1 tyrosine phosphorylation was strongly enhanced in livers of obese *Jnk1*^{-/-} mice in comparison with obese *Jnk1*^{+/+} controls (Fig. 4e and f). We also observed improvement in insulin-induced phosphorylation of the 95-kDa β subunit of the insulin receptor (Fig. 4e and f). However, the increase in IRS-1 tyrosine phosphorylation was more marked and was consistent with decreased Ser 307 phosphorylation. These results demonstrate that the absence of JNK1 enhances the signalling capacity of the insulin receptor, at least in part, through its effects on IRS-1 phosphorylation. It is, of course, possible that additional mechanisms might also be involved in JNK action or the link of serine phosphorylated IRS-1 to insulin resistance.

Nevertheless, this study provides evidence that obesity is associated with abnormally elevated JNK activity, predominantly provided by JNK1. Importantly, the absence of JNK1 results in substantial protection from obesity-induced insulin resistance. Abnormal production of inflammatory cytokines such as TNF- α (ref. 3) and increased concentrations of FFAs⁴ are crucial players in obesity-induced insulin resistance. Induction of insulin resistance by these mediators involves inhibitory serine phosphorylation of IRS-1 (refs 21, 22). Both TNF- α and FFAs are potent activators of JNK^{4,6}, which in turn phosphorylates IRS-1 at Ser 307 (ref. 2). Our studies provide strong evidence that JNK1 is indeed a crucial component of the biochemical pathway responsible for obesity-induced insulin resistance *in vivo*. There is also genetic evidence demonstrating that increased JNK activity caused by loss-of-function mutations in the JNK scaffold protein JIP1 is causal to type 2 diabetes in humans²³. We therefore suggest that a selective interference with JNK1 activity presents an attractive opportunity for the treatment of human obesity, insulin resistance and type 2 diabetes, the most prevalent metabolic diseases worldwide. □

Methods

Generation of mice deficient in JNK1 and JNK2

Generation of *Jnk1*^{-/-} and *Jnk2*^{-/-} mice was as described^{18,19}. All experimental mice were backcrossed five generations to C57Bl/6J and generated from intercrosses between these double heterozygotes and groups were derived from littermates. *ob/ob-Jnk1*^{+/+} and *ob/ob-Jnk1*^{-/-} mice were generated by intercrossing *Jnk1*^{-/-} and *OB/ob* (C57Bl/6J from our in-house colony at Harvard) animals to generate double heterozygotes and subsequent crosses with *OB/ob* breeders to create double homozygous mutant mice.

Diet study and metabolic measurements

Male mice of different genotypes were housed in a barrier-free facility and placed on a high-fat/high-carbohydrate diet *ad libitum* (Diet F3282; Bioserve, Frenchtown, NJ, USA) at 4 weeks of age and were followed for a period of 12 weeks⁶. Biochemical analyses and tolerance tests were performed as described⁶. The polyclonal rabbit anti-mouse ACRP30/

adiponectin antibody was a gift from P. Scherer and was used as described¹³. The polyclonal rabbit anti-mouse resistin antibody was a gift from Affinity Bioreagents Inc..

Measurement of JNK activity and protein concentrations

Tissue extracts (600 μ g of protein) were mixed with 20 μ l of glutathione S-transferase (GST)-agarose resin suspension (Sigma) to which 5 μ g of GST-c-Jun (1–79) were bound. The mixture was agitated at 4 °C overnight, pelleted by centrifugation and washed twice; JNK activity was measured as described²⁴.

Measurement of insulin receptor and IRS-1 phosphorylation *in vivo*

After an overnight fast, mice were anaesthetized as described⁶ and 25 mIU kg⁻¹ insulin (Eli Lilly) or an equal volume of vehicle was administered through the portal vein. Tissues were collected in liquid nitrogen 120 s after injection. Serine phosphorylation of IRS-1 was studied in livers collected from mice without any treatment. Protein extracts (1 mg) from tissue samples were prepared and analysed as described⁶. Antibodies were purchased from Santa Cruz (anti-insulin-receptor- β and anti-phosphotyrosine) or Upstate Biotechnology (anti-IRS-1 and anti-IRS-1-pSer307).

Received 26 April; accepted 3 September 2002; doi:10.1038/nature01137.

1. Saltiel, A. R. & Kahn, C. R. Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* **414**, 799–806 (2001).
2. Aguirre, V., Uchida, T., Yenush, L., Davis, R. & White, M. F. The c-Jun NH₂-terminal kinase promotes insulin resistance during association with insulin receptor substrate-1 and phosphorylation of Ser³⁰⁷. *J. Biol. Chem.* **275**, 9047–9054 (2000).
3. Sethi, J. K. & Hotamisligil, G. S. The role of TNF alpha in adipocyte metabolism. *Semin. Cell. Dev. Biol.* **10**, 19–29 (1999).
4. Boden, G. Role of fatty acids in the pathogenesis of insulin resistance and NIDDM. *Diabetes* **46**, 3–10 (1997).
5. Must, A. *et al.* The disease burden associated with overweight and obesity. *J. Am. Med. Assoc.* **282**, 1523–1529 (1999).
6. Uysal, K. T., Wiesbrock, S. M., Marino, M. W. & Hotamisligil, G. S. Protection from obesity-induced insulin resistance in mice lacking TNF- α function. *Nature* **389**, 610–614 (1997).
7. Hotamisligil, G. S. & Spiegelman, B. M. *Diabetes Mellitus* (eds LeRoith, D., Taylor, S. I. & Olefsky, J. M.) 651–658 (Lippincott Williams & Wilkins, Philadelphia, 2000).
8. Chang, L. & Karin, M. Mammalian MAP kinase signalling cascades. *Nature* **410**, 37–40 (2001).
9. Shin, E. A. *et al.* Arachidonic acid induces the activation of the stress-activated protein kinase, membrane ruffling and H₂O₂ production via a small GTPase Rac1. *FEBS Lett.* **452**, 355–359 (1999).
10. Rizzo, M. T., Leaver, A. H., Yu, W. M. & Kovacs, R. J. Arachidonic acid induces mobilization of calcium stores and c-jun gene expression: evidence that intracellular calcium release is associated with c-jun activation. *Prostaglandins Leukot. Essent. Fatty Acids* **60**, 187–198 (1999).
11. Kyriakis, J. M. & Avruch, J. Sounding the alarm: protein kinase cascades activated by stress and inflammation. *J. Biol. Chem.* **271**, 24313–24316 (1996).
12. Yamauchi, T. *et al.* The fat-derived hormone adiponectin reverses insulin resistance associated with both lipodystrophy and obesity. *Nature Med.* **7**, 941–946 (2001).
13. Berg, A. H., Combs, T. P., Du, X., Brownlee, M. & Scherer, P. E. The adipocyte-secreted protein Acrp30 enhances hepatic insulin action. *Nature Med.* **7**, 947–953 (2001).
14. Stepan, C. M. *et al.* The hormone resistin links obesity to diabetes. *Nature* **409**, 307–312 (2001).
15. Kim, K. H., Lee, K., Moon, Y. S. & Sul, H. S. A cysteine-rich adipose tissue-specific secretory factor inhibits adipocyte differentiation. *J. Biol. Chem.* **276**, 11252–11256 (2001).
16. Davis, R. J. Signal transduction by the JNK group of MAP kinases. *Cell* **103**, 239–252 (2000).
17. Dong, C. *et al.* Defective T cell differentiation in the absence of Jnk1. *Science* **282**, 2092–2095 (1998).
18. Sabapathy, K. *et al.* JNK2 is required for efficient T-cell activation and apoptosis but not for normal lymphocyte development. *Curr. Biol.* **9**, 116–125 (1999).
19. Sabapathy, K. *et al.* c-Jun NH₂-terminal kinase (JNK)1 and JNK2 have similar and stage-dependent roles in regulating T cell apoptosis and proliferation. *J. Exp. Med.* **193**, 317–328 (2001).
20. Shulman, G. I. Cellular mechanisms of insulin resistance. *J. Clin. Invest.* **106**, 171–176 (2000).
21. Hotamisligil, G. H. *et al.* IRS-1-mediated inhibition of insulin receptor tyrosine kinase activity in TNF- α - and obesity-induced insulin resistance. *Science* **271**, 665–668 (1996).
22. Griffin, M. E. *et al.* Free fatty acid-induced insulin resistance is associated with activation of protein kinase C θ and alterations in the insulin signaling cascade. *Diabetes* **48**, 1270–1274 (1999).
23. Waechter, G. *et al.* The gene MAPK8IP1, encoding islet-brain-1, is a candidate for type 2 diabetes. *Nature Genet.* **24**, 291–295 (2000).
24. Hibi, M., Lin, A., Smal, T., Minden, A. & Karin, M. Identification of an oncoprotein- and UV-responsive protein kinase that binds and potentiates the c-Jun activation domain. *Genes Dev.* **7**, 2135–2148 (1993).
25. Uysal, K. T., Scheja, L., Wiesbrock, S. M., Bonner-Weir, S. & Hotamisligil, G. S. Improved glucose and lipid metabolism in genetically obese mice lacking aP2. *Endocrinology* **141**, 3388–3396 (2000).
26. Bennett, B. L. *et al.* SP600125, an anthrapyrazolone inhibitor of Jun N-terminal kinase. *Proc. Natl Acad. Sci. USA* **98**, 13681–13686 (2001).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements This work is in part supported by grants from the National Institutes of Health (M.K. and G.S.H.) and the Pew Foundation (G.S.H.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.S.H. (e-mail: ghotamis@hsph.harvard.edu).

Arabidopsis boron transporter for xylem loading

Junpei Takano*, Kyotaro Noguchi*†, Miho Yasumori*, Masaharu Kobayashi*, Zofia Gajdos*, Kyoko Miwa*, Hiroaki Hayashi*, Tadakatsu Yoneyama* & Toru Fujiwara*

* Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

Boron deficiency hampers the productivity of 132 crops in more than 80 countries¹. Boron is essential in higher plants primarily for maintaining the integrity of cell walls^{2–4} and is also beneficial and might be essential in animals⁵ and in yeast⁶. Understanding the molecular mechanism(s) of boron transport is crucial for alleviating boron deficiency. Here we describe the molecular identification of boron transporters in biological systems. The *Arabidopsis thaliana* mutant *bor1-1* is sensitive to boron deficiency^{7,8}. Uptake studies indicated that xylem loading is the key step for boron accumulation in shoots with a low external boron supply and that the *bor1-1* mutant is defective in this process. Positional cloning identified BOR1 as a membrane protein with homology to bicarbonate transporters in animals. Moreover, a fusion protein of BOR1 and green fluorescent protein (GFP) localized to the plasma membrane in transformed cells. The promoter of *BOR1* drove GFP expression in root pericycle cells. When expressed in yeast, BOR1 decreased boron concentrations in cells. We show here that BOR1 is an efflux-type boron transporter for xylem loading and is essential for protecting shoots from boron deficiency.

The mechanism of boron membrane transport in plants has long been controversial: the most widely accepted theory was that of a passive transport of boric acid, whereas recent physiological studies indicated the involvement of active transport^{2,3,9,10}.

The *Arabidopsis thaliana* mutant *bor1-1* requires high levels of boron for normal growth⁷. Growth impairment of *bor1-1* mutant plants under low boron supply is more evident in leaves than in roots⁸ and the mutation carried by *bor1-1* mutant plants is suggested to affect the root-to-shoot transport of boron in plants, rather than the structural function of boron in cell walls¹¹. Here we describe the identification of BOR1 as a boron transporter essential for alleviating boron deficiency.

The concentration of boron in rosette leaves of wild-type and *bor1-1* mutant plants increased with the concentration of boric acid in the nutrient solution (Fig. 1a). Boron concentrations in the rosette leaves of wild-type Col-0 plants were always higher than those in the *bor1-1* plants under the nutrient conditions tested, indicating that *bor1-1* plants have a reduced capacity to accumulate boron in leaves over a wide range of external boron concentrations.

The pathway of nutrient transport from the root surface to the shoot includes at least two transmembrane transport events: import into epidermal or cortex cells and export from pericycle or xylem parenchyma cells into the stelar apoplasm (xylem loading). As has been shown for phosphate¹² and potassium¹³, xylem loading is a key step for nutrient accumulation in the shoots.

The transport events of boron in *Arabidopsis* plants were studied with ¹⁰B-enriched boron as a tracer. We examined concentrations of tracer boron in root cell saps, representing the water-soluble fraction in roots^{11,14}, and in xylem exudates after tracer boron had been supplied for 24 h. The concentration of tracer boron in root cell saps increased in proportion to tracer boron concentration in the medium in both genotypes (Fig. 1b), suggesting that uptake into roots occurs mainly by passive transport. The concentration of

tracer boron in xylem exudates of the *bor1-1* plants also followed a linear concentration dependence, whereas in the wild-type plants a combination of saturable and linear concentration dependence was observed (Fig. 1c). When the tracer boron supply was 30 μ M or lower, tracer boron concentrations in xylem exudates of wild-type plants were significantly higher than those in the medium, the root cell sap of the wild type plants and the xylem exudate of *bor1-1* plants. These results indicate that xylem loading is the key step for boron accumulation in shoots at low external boron supply and that BOR1 is an essential component in this process.

The *BOR1* locus is located on the lower arm of chromosome 2 (ref. 7). Analysis of 1,038 F₂ plants delimited *BOR1* to a 15.1-kilobase (kb) region between newly generated molecular markers at positions 19,383 kb and 19,399 kb. Nucleotide sequences of this region in the genome of *bor1-1* and *bor1-2* mutants (ethylmethane sulphonate mutants)¹⁵ revealed that each mutant contains a different single base substitution in the hypothetical open reading frame (ORF) At2g47160 (Fig. 2a), each causing a different amino-acid substitution in the predicted protein. A genomic DNA fragment containing wild-type At2g47160 was then introduced into the *bor1-1* mutant and several independent transgenic plants were generated. Growth of the transgenic *bor1-1* plants was indistinguishable from that of the Col-0 wild type on low-boron medium. In contrast, growth of the non-transgenic *bor1-1* plants was inhibited on the same medium (Fig. 2b), establishing that At2g47160 corresponds to the *BOR1* gene.

The *BOR1* ORF (Genbank accession no. AB073713) was determined by sequencing a corresponding cDNA, RZ119b07, obtained from Kazusa DNA Research Institute (Chiba, Japan). Comparison between the cDNA and genomic sequences revealed that *BOR1* has 12 exons (Fig. 2a). On the basis of the nucleotide sequence, *BOR1* was predicted to encode a polypeptide of 704 amino acids containing 10 putative transmembrane domains (Fig. 2c). The mutations found in the *bor1-1* and *bor1-2* alleles were located within the second transmembrane domain.

Database searches revealed that there are six similar predicted proteins to BOR1 in *Arabidopsis*: At1g15460, At1g74810, At3g06450, At3g62270, At4g32510 and At5g25430. The amino acid sequences of the predicted proteins were 52–90% identical to BOR1 (Fig. 2d). BOR1 also exhibited strong similarity to expressed sequence tag (EST) clones from diverse plant species, including both angiosperms (monocots and dicots) and gymnosperms (Fig. 2d), indicating that BOR1 belongs to a group of highly conserved membrane proteins in plants. BOR1 is also similar to bicarbonate transporters of animal origins, including Cl[−]/HCO₃[−] exchangers (anion exchangers), Na⁺–HCO₃[−] co-transporters and Na⁺-driven Cl[−]/HCO₃[−] exchangers^{16,17} (Fig. 2d and Supplementary Fig. 1). BOR1 and bicarbonate transporters of animals also show

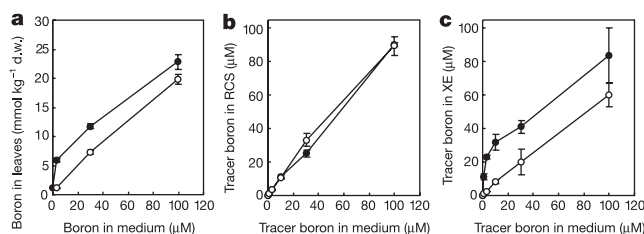


Figure 1 Accumulation of boron in *Arabidopsis* mutant *bor1-1*. **a**, Boron concentrations in rosette leaves. Wild-type (filled circles) and *bor1-1* mutant (open circles) plants were grown hydroponically for 4 weeks supplied with 0.3 (only for wild type), 3, 30 or 100 μ M boric acid, as described previously⁸. **b, c**, Accumulation of boron in root-cell sap (RCS, **b**) and xylem exudate (XE, **c**). The plants were grown for 24 or 23 days with 30 μ M boron and supplied for 24 h with 1, 3, 10, 30 or 100 μ M tracer boron. Tracer boron concentrations in root cell sap and xylem exudate were then determined. All data are means $\pm$ s.d. ($n = 3$ or 4). d.w., dry weight.

† Present address: Forestry and Forest Products Research Institute, Ibaraki 305-8687, Japan.

significant similarity to a membrane protein of yeast, YNL275w (Fig. 2c and Supplementary Fig. 1). The amino acid sequence of YNL275w was 27% identical to BOR1 in a 596-amino-acid overlap by FASTA analysis¹⁸. Subcellular fractionation and immunofluorescence labelling experiments indicated that YNL275w was localized to the plasma membrane, and a binding assay suggested that HCO_3^- , I^- , Br^- , NO_3^- and Cl^- were the potential substrates¹⁹. Both mutations in *bor1-1* and *bor1-2* alleles alter amino-acid residues that are highly conserved among the similar proteins in plants, yeast and animals (Fig. 2d).

To investigate the subcellular localization of BOR1, a translational fusion between BOR1 and green fluorescent protein (GFP)²⁰ was delivered into tobacco leaf cells by particle bombardment. Cells expressing the BOR1–GFP fusion showed a GFP signal only at the periphery of the cell (Fig. 3a), whereas in cells expressing free GFP the signal was also observed in a region around the nucleus and in the cytosol as evidenced by cytoplasmic strands (Fig. 3b). Together with the presence of transmembrane domains, these results suggest that BOR1 is a plasma-membrane-localized protein.

Cell-type specificity of BOR1 expression was investigated with the T2 generation of transgenic *Arabidopsis* plants carrying the ORF for

GFP under the control of the BOR1 promoter region (2.9 kb). The GFP fluorescence signal was observed in roots (Fig. 3c) and was localized in the pericycle, located at the outmost layer of the stele and the inside of the endodermis (Fig. 3d). Pericycle-specific expression was observed in five independent transgenic lines analysed. The GFP fluorescence signal could not be clearly observed in shoots although the transcripts of BOR1 were detected from both shoots and roots of wild-type plants by polymerase chain reaction with reverse transcription (Supplementary Fig. 2). The *Arabidopsis* stelar K^+ outward rectifying channel (SKOR) is responsible for xylem loading of K^+ and is also expressed in the root pericycle¹⁴. Pericycle-specific expression of the BOR1 promoter supports the hypothesis that BOR1 is involved in the xylem loading of boron.

We examined the boron transport activity of BOR1 in the yeast *Saccharomyces cerevisiae* by first comparing boron concentrations between a *S. cerevisiae* mutant carrying an insertion mutation in YNL275w and the corresponding wild-type strain (Fig. 4a). Cells were first grown to mid-exponential phase in liquid medium without boron supplementation and then incubated for 60 min in the presence of 500 μM boric acid, followed by measurements of soluble and bound boron concentrations. Surprisingly, concentrations of

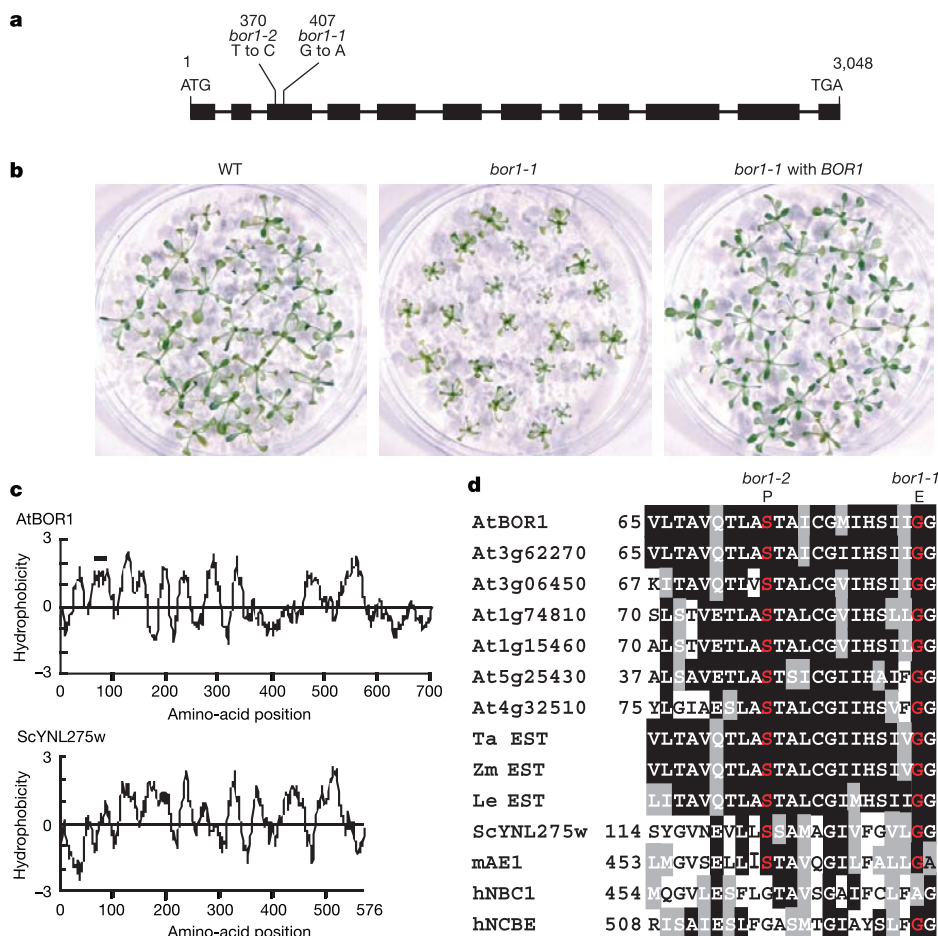


Figure 2 Identification and sequence analysis of BOR1. **a**, Structure of the BOR1 gene, including positions and nucleotide changes of the *bor1-1* and *bor1-2* mutations. Nucleotide positions are relative to the start codon. Filled boxes indicate the ORF and lines between boxes indicate introns. **b**, Complementation of *bor1-1* by the wild-type BOR1 gene. Wild-type plants (left), *bor1-1* mutant plants (centre), and transgenic *bor1-1* plants (T3 generation, homozygous for T-DNA insertion) containing the wild-type BOR1 gene including its own promoter (right) were grown for 16 days on medium containing 3 μM boric acid. **c**, Hydrophobicity plots of BOR1 and *Saccharomyces cerevisiae* YNL275w. Plots were generated as described by Kyte and Doolittle³⁰. The running window was 17 amino acids. The second putative transmembrane domain is marked with a blue bar.

d, Predicted second transmembrane domain of BOR1 aligned with similar domains from predicted proteins of *Arabidopsis* and related sequences of other organisms. GenBank accession numbers for the EST clones are BG605038 (Ta, *Triticum aestivum*), AW330878 (Zm, *Zea mays*), and AW031603 (Le, *Lycopersicon esculentum*). The SwissProt accession numbers for mouse anion exchanger 1 (mAE1), human NCB1 and NCBE are P04919, NP003750 and BAB18301, respectively. The transmembrane domains of BOR1 were deduced by the program TMPRED available at <http://www.ch.embnet.org>. The most conserved amino-acid residues at each position among the proteins are highlighted in black and conservative substitutions in grey. Red characters indicate conserved residues that were replaced in *bor1-2* (S to P) and *bor1-1* (G to E) genomes.

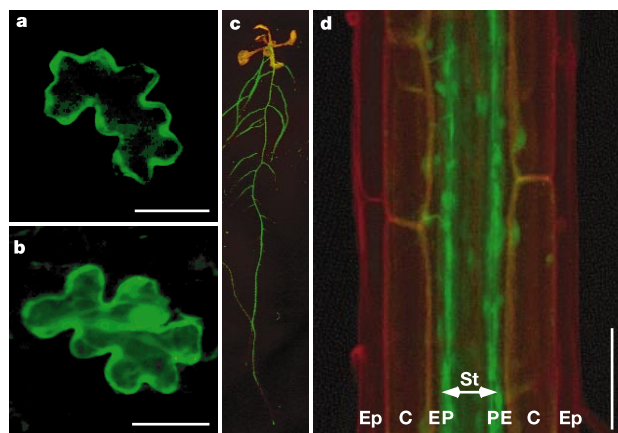


Figure 3 Subcellular localization of BOR1 and cell-type-specific localization of *BOR1* expression. **a, b**, Subcellular localization of BOR1–GFP fusion (**a**) and GFP (**b**) proteins. Tobacco leaf cells were bombarded with expression plasmids containing genes transcribed from a 35S promoter. **c, d**, Cell-type-specific localization of GFP in transgenic *Arabidopsis* expressing the *BOR1*–promoter–GFP fusion gene. An image of the green GFP fluorescence signals in whole plants (**c**) and a laser scanning confocal microscopic image of the green GFP fluorescence signals in mature portion of roots (**d**) are shown. The images were merged with the red signals from chlorophyll (**c**) and propidium-iodide-stained cell wall (**d**). Ep, epidermis; C, cortex; E, endodermis; P, pericycle; St, stele. Scale bar, 50 µm.

soluble boron were about 13-fold higher in the *ynl275w* mutant than in the wild type, whereas bound boron concentrations were similar between the two strains. We then expressed the *BOR1* gene under control of the *GAL1* promoter in the *ynl275w* cells. The soluble boron concentrations in the cells expressing the *BOR1* were about one-third those in the cells carrying the empty vector, whereas the bound boron concentrations were similar between these two yeast lines (Fig. 4a). At physiological pH, boron exists mainly as undissociated boric acid ($pK_a = 9.25$). Considering the relatively high permeability of the plasma membrane for boric acid^{10,21–23}, which allows the passive diffusion of boric acid into cells, these results suggest that BOR1 and YNL275w have boron-efflux trans-

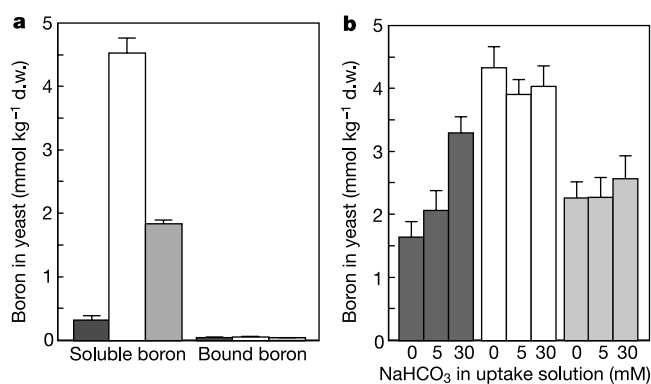


Figure 4 Concentrations of boron in yeast cells. **a**, Boron concentration in wild-type cells transformed with pYES2 (black columns), *ynl275w* mutant cells transformed with pYES2 (white columns) or pYES2 with *BOR1* cDNA (grey columns). Cells were grown to mid-exponential phase in liquid medium containing galactose²⁹ and transferred to the same liquid medium (pH 5.5) containing 500 µM boric acid. Cultures were then incubated for 60 min, harvested by centrifugation, washed twice with distilled water and boiled with water for 30 min. Soluble boron represents boron in the centrifuged supernatant, and bound boron represents boron in the pellet. **b**, Effects of NaHCO₃ on boron concentration in yeast. Boron transport experiments were performed as above except that the media used for exposure of boron contained 0, 5 or 30 mM NaHCO₃ at pH 7. Boron concentrations in the soluble fraction are shown. Data are means $\pm$ s.d. for four or five independent experiments. d.w., dry weight.

port activity. Although both BOR1 and YNL275w are similar to bicarbonate transporters on a sequence basis (Fig. 2d), the boron transport activity of BOR1, in contrast to YNL275w, was little affected by NaHCO₃ in the uptake solution (Fig. 4b).

The present results establish that BOR1 is a boron transporter mediating boron export from pericycle cells into the root stelar apoplasm against a concentration gradient. Consequently, BOR1 increases the concentrations of boron in xylem sap and shoot organs, thereby acting to protect the shoot from boron deficiency. To our knowledge, BOR1 and YNL275w represent the first identification of boron transporters in biological systems and provide a molecular basis for active boron transport of boron in higher plants and yeast. □

Methods

Positional identification of *BOR1*

1,038 F₂ plants from the *bor1-1* mutant (Col-0 background) and *Ler* crosses⁷ were scored for the *Bor1*⁻ phenotype (growth retardation at 3 µM boric acid) and molecular markers including cleaved amplified polymorphic sequences markers at positions 19,383 kb and 19,399 kb on chromosome 2 (see Supplementary Information 3). For complementation *in vivo*, a genomic *Sall*–*BglII* fragment spanning from 1,158 base pairs (bp) upstream of the *BOR1* initiation codon to 418 bp downstream of the stop codon was cloned into pBI101.1 (ref. 24) (replacing the *HindIII*–*EcoRI* fragment). Using *Agrobacterium tumefaciens* C58C1 Rif^R (pGV2260)²⁵ the construct was introduced into *Arabidopsis bor1-1* mutant plants²⁶. Growth of the resulting kanamycin-resistant T2 transgenic plants homozygous for the T-DNA insertion were tested on solid MGR1 medium²⁷ containing 2% (w/v) sucrose and 0.2% (w/v) Gellan gum (Wako Pure Chemicals) with 3 µM boric acid.

Transport measurements in *A. thaliana*

Seeds were sown on black cotton cloth fixed on the top of 50-ml plastic tubes. During the first 2 weeks of cultivation, the MGR1 medium (pH 5.8, 18.9 mOsm)²⁷ were replenished every 2–3 days and thereafter, the solutions were exchanged every 4–5 days. Plants grown for 23 or 24 days with 30 µM boric acid of natural abundance (¹⁰B:¹¹B = 20:80) were transferred to medium containing ¹⁰B-enriched boric acid (¹⁰B:¹¹B = 96:4). After 24 h of application of tracer boric acid, shoots were cut off at their hypocotyls, xylem exudates were collected from the cut ends of the roots for 2 h and roots were harvested and frozen. Frozen samples were thawed, transferred to 1.5-ml centrifugal filter tubes (pore size 0.2 µm, Ultrafree-MC; Millipore) and centrifuged at 10,000g at 4 °C for 10 min. Filtered solutions were taken as root cell saps.

Imaging of GFP expression

BOR1–sGFP fusion constructs were generated as described in the Supplementary Information 4. The resulting plasmid (pTF458) carried a BOR1–GFP fusion gene under control of the cauliflower mosaic virus (CaMV) 35S RNA promoter.

Gold particles (1 µm diameter) were coated with pTF458 or pTH2 (ref. 20), carrying CaMV 35S RNA–promoter–GFP, and bombarded into tobacco (*Nicotiana tabacum* cv. Xanthi sx) leaf cells with a helium-gas-driven particle accelerator (PDS-1000/He; Bio-Rad) operated at a pressure of 1,100 lb in⁻² (7.6 MPa) and a chamber vacuum of 26 inches of mercury. Leaves were kept at 25 °C in darkness for 24 h and fluorescence was observed by fluorescence microscopy (BX-FLA; Olympus) equipped with a chilled, colour, charge-coupled device camera (model 3204 C; Olympus).

Procedures for the construction of the *BOR1* promoter–GFP fusion are described in Supplementary Information 5. Transformation of *Arabidopsis* Col-0 was performed as described above. Transgenic plants were grown on solid MGR1 medium²⁷ containing 2% (w/v) sucrose and 1.5% (w/v) Gellan gum with 3 µM boric acid. Imaging of GFP in whole plants was performed with a Fluoroimager 595 and ImageQuant software (Molecular Dynamics). Plants were scanned with an Ar laser (488 nm) and the green fluorescence of GFP (515–545 nm) and autofluorescence of chlorophyll (610 nm) were measured. Laser scanning confocal microscopy of GFP expression in the roots treated with 10 µg ml⁻¹ propidium iodide (Molecular Probes) was performed with a Zeiss LSM 510. Dual-channel imaging was performed with a 488-nm excitation wavelength and emission windows of 505–530 nm for GFP and 585 nm for propidium iodide. Images were merged using Adobe Photoshop 4.0J software (Adobe Systems).

Transport measurements in yeast cells

Strains of *S. cerevisiae* used in this study were BY4741 (*MATa his2DO met15DO ura3DO*) and 1169. Strain 1169 was constructed from BY4741 by insertional mutagenesis of the YNL275w ORF²⁸. BY4741 and 1169 cells were transformed with either pYES2 (Invitrogen) or pYES2 with a *BOR1* cDNA insert. Unless otherwise mentioned, yeast cells were grown in medium containing constituents of the synthetic minimal medium²⁹ supplemented with 2% glucose, 20 mg l⁻¹ His, 30 mg l⁻¹ Leu and 20 mg l⁻¹ Met. For induction of the *GAL1* promoter, glucose was replaced with galactose in the medium. For measurements of boron concentration, cells at mid-exponential phase were harvested and transferred to the same medium containing 500 µM boric acid and 50 mM MES adjusted to pH 5.5 with Tris. For the NaHCO₃ study, the medium was adjusted to pH 7 with Tris. Cells were incubated at 30 °C with shaking for 60 min, harvested by centrifugation and washed twice with deionized water (MilliQ; Millipore) then boiled for 30 min with deionized water. All centrifugation steps were performed at 3,000g for 2 min at 4 °C. Clarified supernatants and pellets were taken separately for determination of intracellular free and bound boron,

respectively. In media with glucose (*GAL1* promoter inactive), both soluble and bound boron concentrations were similar between yeast lines (data not shown).

Analysis of boron concentration

Determination of boron concentration by inductively coupled plasma mass spectrometry were performed as described previously^{7,8,11} with some modifications (see Supplementary Information 6). Tracer boron concentrations were calculated as described previously⁷.

Received 28 January; accepted 3 September 2002; doi:10.1038/nature01139.

- Shorrocks, V. M. The occurrence and correction of boron deficiency. *Plant Soil* **193**, 121–148 (1997).
- Dannel, F., Pfeffer, H. & Römhild, V. Update on boron in higher plants – uptake, primary translocation and compartmentation. *Plant Biol.* **4**, 193–204 (2002).
- Brown, P. H. et al. Boron in plant biology. *Plant Biol.* **4**, 205–223 (2002).
- O'Neill, M. A., Eberhard, S., Albersheim, P. & Darvill, A. G. Requirement of borate cross-linking of cell wall rhamnogalacturonan II for *Arabidopsis* growth. *Science* **294**, 846–849 (2001).
- Nielsen, E. H. The emergence of boron as nutritionally important throughout the life cycle. *Nutrition* **16**, 512–514 (2000).
- Bennett, A., Rowe, R. L., Soch, N. & Eckhart, C. D. Boron stimulates yeast (*Saccharomyces cerevisiae*) growth. *J. Nutr.* **129**, 2236–2238 (1999).
- Noguchi, K. et al. *bor1-1*, an *Arabidopsis thaliana* mutant that requires a high level of boron. *Plant Physiol.* **115**, 901–906 (1997).
- Takano, J., Yamagami, M., Noguchi, K., Hayashi, H. & Fujiwara, T. Preferential translocation of boron to young leaves in *Arabidopsis thaliana* regulated by the *BOR1* gene. *Soil Sci. Plant Nutr.* **47**, 345–357 (2001).
- Dannel, F., Pfeffer, H. & Römhild, V. Characterization of root boron pools, boron uptake and boron translocation in sunflower using the stable isotopes ¹⁰B and ¹¹B. *Aust. J. Plant Physiol.* **27**, 397–405 (2000).
- Stangoulis, J. C. R., Reid, R. J., Brown, P. H. & Graham, R. D. Kinetic analysis of boron transport in *Chara*. *Planta* **213**, 142–146 (2001).
- Noguchi, K. et al. Defect in root-shoot translocation of boron in *Arabidopsis thaliana* mutant *bor1-1*. *J. Plant Physiol.* **156**, 751–755 (2000).
- Poirier, Y., Thomas, S., Somerville, C. & Schiefelbein, J. A mutant of *Arabidopsis* deficient in xylem loading of phosphate. *Plant Physiol.* **97**, 1087–1093 (1991).
- Gaymard, F. et al. Identification and disruption of a plant shaker-like outward channel involved in K⁺ release into the xylem sap. *Cell* **94**, 647–655 (1998).
- Dannel, F., Pfeffer, H. & Römhild, V. Compartmentation of boron in roots and leaves of sunflower as affected by boron supply. *J. Plant Physiol.* **153**, 615–622 (1998).
- Yasumori, M., et al. *Plant Nutrition: Molecular Biology and Genetics* (eds Nielsen, E. H. & Jensen, A.) 269–275 (Kluwer, Dordrecht, 1999).
- Alper, S. L., Darman, R. B., Chernova, M. N. & Dahl, N. K. The AE gene family of Cl[−]/HCO₃[−] exchangers. *J. Nephrol.* **15** suppl. 5, S41–S53 (2002).
- Soleimani, M. Na⁺/HCO₃[−] cotransporters (NBC): Expression and regulation in the kidney. *J. Nephrol.* **15** suppl. 5, S32–S40 (2002).
- Pearson, W. R. & Lipman, D. J. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448 (1988).
- Zhao, R. & Reithmeier, R. A. F. Expression and characterization of the anion transporter homologue YNL275w in *Saccharomyces cerevisiae*. *Am. J. Physiol. Cell Physiol.* **281**, C33–C45 (2001).
- Chiu, W. L. et al. Engineered GFP as a vital reporter in plants. *Curr. Biol.* **6**, 325–330 (1996).
- Raven, J. A. Short- and long-distance transport of boric acid in plants. *New Phytol.* **84**, 231–249 (1980).
- Dordas, C., Chrispeels, M. J. & Brown, P. H. Permeability and channel-mediated transport of boric acid across membrane vesicles isolated from squash roots. *Plant Physiol.* **124**, 1349–1361 (2000).
- Dordas, C. & Brown, P. H. Permeability and the mechanism of transport of boric acid across the plasma membrane of *Xenopus laevis* oocytes. *Biol. Trace Elem. Res.* **81**, 127–139 (2001).
- Jefferson, R. A., Kavanagh, T. A. & Bevan, M. W. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901–3907 (1987).
- Deblaere, R. et al. Efficient octopine Ti plasmid-derived vectors for Agrobacterium-mediated gene transfer to plants. *Nucleic Acids Res.* **13**, 4777–4788 (1985).
- Bechtold, N., Ellis, J. & Pelletier, G. In planta Agrobacterium mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C.R. Acad. Sci. Paris Life Sci.* **316**, 1194–1199 (1993).
- Fujiwara, T., Hirai, M. Y., Chino, M., Komeda, Y. & Naito, S. Effects of sulfur nutrition on expression of the soybean seed storage protein genes in transgenic petunia. *Plant Physiol.* **99**, 263–268 (1992).
- Winzler, E. A. et al. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906 (1999).
- Sherman, F. Getting started with yeast. *Method. Enzymol.* **194**, 3–21 (1991).
- Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com/nature>).

Acknowledgements We thank S. Naito, N. von Wirén and P. Walch-Liu for critical reading of the manuscript; Y. Niwa for providing the sGFP construct; K. Mukaikawato, M. Hashimoto and M. Sugawara for technical contributions; and F. Kurisu for use of the facilities of the Research Center for Water Environment Technology in the School of Engineering, the University of Tokyo. RZ119b07 was kindly provided by Kazusa DNA Research Institute; BAC T3D7 clone was kindly provided by the *Arabidopsis* Biological Resource Center at Ohio State University. This work is supported in part by grants to T.F. from the Japanese Ministry of Education, Sports, Culture, Science, and Technology and the Graduate School of Agricultural Life Sciences, the University of Tokyo.

Competing interests statement The authors declare competing financial interests: details accompany the paper on Nature's website (<http://www.nature.com>).

Correspondence and requests for materials should be addressed to T.F. (e-mail: atorufu@mail.ecc.u-tokyo.ac.jp) Genbank accession number of BOR1 is AB073713.

In silico simulations reveal that replicators with limited dispersal evolve towards higher efficiency and fidelity

Péter Szabó*†, István Scheuring*, Tamás Czárán* & Eörs Szathmáry*‡

* Department of Plant Taxonomy and Ecology, Research Group of Ecology and Theoretical Biology, Eötvös University, 1/c Pázmány P. sétány, H-1117 Budapest
† Adaptive Dynamics Network, International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria

‡ Collegium Budapest (Institute for Advanced Study) 2 Szentháromság u., H-1014 Budapest, Hungary

The emergence of functional replicases, acting quickly and with high accuracy, was crucial to the origin of life¹. Although where the first RNA molecules came from is still unknown, it is nevertheless assumed that catalytic RNA enzymes (ribozymes) with replicase function emerged at some early stage of evolution¹. The fidelity of copying is especially important because the mutation load limits the length of replicating templates that can be maintained by natural selection². An increase in template length is disadvantageous for a fixed digit copying fidelity, however, longer molecules are expected to be better replicases. An iteration for longer molecules with better replicase function has been suggested^{3,4} and analysed mathematically⁵. Here we show that more efficient replicases can spread, provided they are adsorbed to a prebiotic mineral surface. A cellular automaton⁶ simulation reveals that copying fidelity, replicase speed and template efficiency all increase with evolution, despite the presence of molecular parasites, essentially because of reciprocal altruism⁷ ('within-species mutualism') on the surface⁸, thus making a gradual improvement of replicase function more plausible.

We considered a population of macromolecules, adsorbed to a surface and built of four different monomers: A, B, C and D. Owing to their catalytic activity, macromolecules located on neighbouring sites of the surface can template-replicate each other, that is, they can build a new macromolecule from free monomers by copying an existing one. In each replication process there are two replicator molecules involved: one is the template, the other acts as a replicase enzyme. We attributed two main properties to replication events: speed and fidelity, which in turn depend on three parameters of the two replicators involved in the process: (1) replicase activity, which expresses how fast the molecule can add a monomer to a new strain while acting as a replicase; (2) replicase fidelity, which measures the accuracy of replication per monomer when the molecule acts as a replicase; (3) template efficiency, which defines the average 'affinity' of the molecule, that is, its tendency to behave as a template against others.

Replication speed depends on the activity of the replicase and the quality of the template (as explained below in detail) such that higher replicase activity and template efficiency result in faster replication. Given two neighbouring replicator molecules, L and M, on the surface, one of two different replication events can occur: either L as a replicase copies M as a template, or vice versa.

For simplicity, we assumed that replicase activity, replicase fidelity and template efficiency depend on the primary structure of the replicator molecule as follows: A, B and C type monomers each affect one of the three relevant replicator properties: A enhances template efficiency, B increases replicase activity, and C is beneficial in terms of replicase fidelity. D is a neutral ('dummy') monomer of no direct effect on replication. We assumed that the template efficiency $t(n_A)$ of a certain replicator molecule is an

increasing sigmoid function of the number n_A of A monomers in the longest contiguous poly-A section within the macromolecule:

$$t(n_A) = \alpha_A + (1 - \alpha_A) \frac{n_A^{\beta_A}}{\gamma_A + n_A^{\beta_A}} \quad \text{if } n_A > 0 \quad (1)$$

$$t(n_A) = 0 \quad \text{if } n_A = 0 \quad (2)$$

With appropriate $\alpha_B, \beta_B, \gamma_B$ and $\alpha_C, \beta_C, \gamma_C$ parameters, replicase activity $r(n_B)$ and replicase fidelity $f(n_C)$ were defined as increasing sigmoid functions of the lengths of contiguous poly-B and poly-C sections in the macromolecule, applying the same mathematical form. Minimal fidelity (at 0 poly-C sequence length) was set to values of 0.7 to 0.9 in different simulations (random monomer insertion itself implies a fidelity of 0.25). Our rationale for choosing a sigmoid function is as follows: (1) There should be saturation after a sufficient number of monomers involved (equivalent to diminishing returns). One reason for this is that the reactions become diffusion-limited, especially on a surface. (2) If a number of critical sites are involved in forming a working active centre, then it is reasonable to assume that there is a synergistic effect between these sites when the molecules are small. The combination of these two effects results in a sigmoid function.

Our modelling of domains that are responsible for the different molecular traits, their possible mutational deterioration and the trade-offs between them is sound. We would expect the same qualitative (although not quantitative) results for the molecularly explicit, but unfeasible, type of simulation.

Even these oversimplified assumptions preserve the general properties of the macromolecule replicators that are most relevant

to our problem. Although we can reliably calculate the two-dimensional structure of RNA molecules, this is not true for three-dimensional structures. Moreover, we do not know how to calculate the three pertinent replicator properties (replicase activity, replicase fidelity and template quality) from an arbitrary, given sequence. Thus we make the assumption that these are associated with separate parts of the molecule, and that they are in a three-way trade-off: in a macromolecule of a given length, one property can be enhanced at expense of the other two. Usually, RNA templates must bear a target structure in order to be accepted by contemporary RNA replicases made of protein¹⁰. Accuracy and speed are analysed in sufficient detail for DNA-dependent DNA polymerases. For example, there are mutator and antimutator mutants of the phage T4 polymerase, and the latter proceed more slowly on the DNA template^{11,12}. In the polymerase of phage RB69, a close relative of phage T4, the polymerase active site and the exonuclease active site are 30 Å apart¹³. It seems that an important determinant of polymerase fidelity is geometric selection of Watson–Crick versus non-Watson–Crick base pairs¹², a result confirmed by recent visualization of DNA polymerase crystals^{14,15}. It has been pointed out that the proposed mechanism could work for ribozyme replicases¹⁶. The three important replicase traits are expected to conflict with one another, especially in the stringent case of small molecules. But this is a conservative assumption, making the emergence of an efficient replicase population more difficult.

During a replication event mutations may occur, that is, the copy may differ from the template. For computational simplicity, base pairing is assumed to be homologous rather than complementary (that is, A pairs with A, C pairs with C, and so on). There are two types of mutation, as known for real replicases or polymerases.

The first type is addition and deletion mutations which occur at a

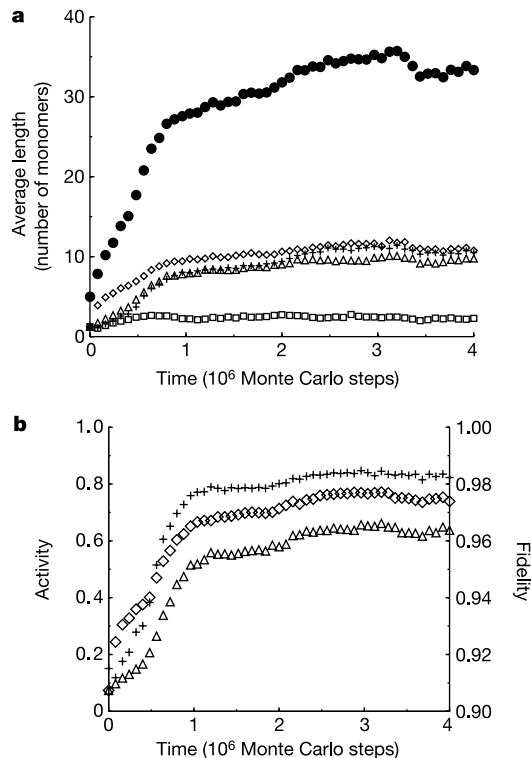


Figure 1 Long-term evolution of replicators (population-wide averages). **a**, Time series of length (closed circles) and of the amount of contributing A (open diamonds), B (open triangles), C (plus signs) and D monomers (open squares); **b**, Evolution of template (open diamonds), replicase (open triangles) catalytic activities and the fidelity of replication (pluses). Catalytic activities are scaled between 0 and 1, see equation (1). Parameter values: $\alpha_A = 0.1$; $\beta_A = 3$; $\gamma_A = 200$; $\alpha_B = 0.1$; $\beta_B = 3$; $\gamma_B = 200$; $\alpha_C = 0.9$; $\beta_C = 2$; $\gamma_C = 5$; $\mu = 6$; P_d (decay rate) = 0.001; $P_{ad} = 0.02$; no diffusion.

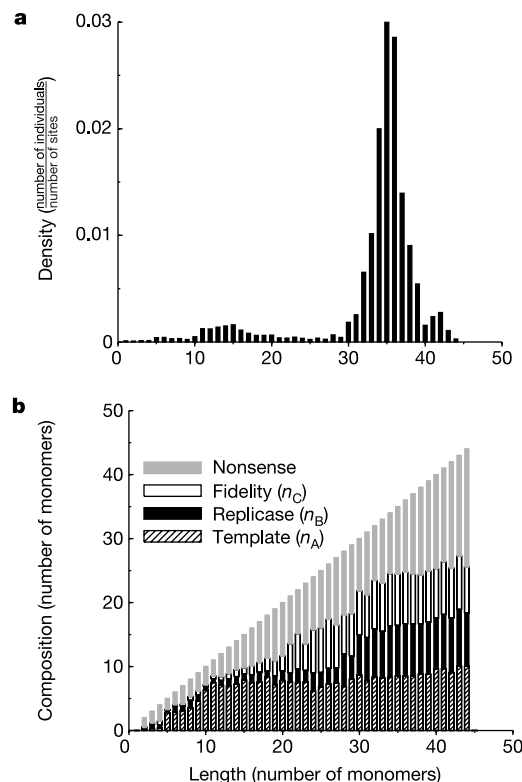


Figure 2 Distribution of replicators at the stationary state. **a**, Density of replicators of different length; **b**, Average length of template (hatched bars), replicase (black boxes), fidelity (white bars) functional sites and of nonsense parts (grey bars) within replicators of different lengths (nonsense parts include D monomers and shorter blocks of A, B and C monomers). (Parameters as before.)

constant probability $P_{ad} = 0.02$. Addition mutation attaches an extra monomer to the end of the molecule under copying; deletion mutation means that the copy breaks at a random position, and one of the fragments is lost, modelling damaging events. These mutations generate variance in the length of replicators.

The second type is that of point (substitution) mutations. Each monomer of the template chain is copied with accuracy equal to the replicase fidelity of the molecule acting as the replicase. If mutation occurs, then one of the three remaining monomer types is inserted into the copy, instead of the proper one.

At $t = 0$ half of the sites are 'inoculated' with five-monomer, random-sequence replicators, and we follow the evolution of the replicator population through many generations. Lineages producing sufficient numbers of accurate copies before decaying persist; others disappear from the selection arena. We show that selection favours the dominant quasispecies that possesses evolutionarily important properties. The quasispecies, thus developed, provides a good starting point for subsequent evolution under

different selection pressures.

Starting from a population of oligomers with very limited functionality, complex, effective replicators emerge and spread (Fig. 1a, b). All the three important replicator properties (replicator activity and fidelity, and template quality) show considerable development. Selection favours replicators performing both replicate and template functions effectively and carrying out replication with high fidelity. The selective advantage of increasing template quality is obvious; a good template is replicated with higher probability by surrounding replicators. However, the selective advantage of replicase activity and of fidelity is a bit surprising. A replicator having these capabilities replicates its neighbours with higher probability, but it needs other efficient replicases around to spread. The advantage of these molecules comes from local interactions resulting in aggregated patterns, thus making reciprocal altruism between good replicases possible.

The monomers enhancing the replicase and template functions increase in number within the molecules, and they organize into functional sites. The evolutionary force shaping functional sites may be estimated by comparing the length of functional groups consisting of A, B and C monomers respectively, to the selectively neutral, contiguous D stretches, which are 5–10 times shorter than the length of the average functional site. The emergence of more complex functional sites results in an overall increase of the average length of replicators.

For selection in favour of faster and more accurate replicases to occur, even the least accurate replicases (those containing no C monomers) must have a fidelity significantly better than 0.25 (representing random monomer insertion), otherwise the system collapses in all cases, owing to the high error rate in replication. This is not a very strict restriction, given the inherent propensity for non-random monomer pairing in all known systems of template replication.

In the course of time this evolution comes to a halt. The diminishing (because of the sigmoid functions involved) advantage of extended functional sites is counterbalanced by the disadvantage of longer replication times due to increased template length. The population reaches a stable distribution of replicators of different lengths and functional structures (Fig. 2a, b). The maximum attainable length depends on the model parameters affecting replication, rather than on nutrient exhaustion. Although there are several sequences that differ in length and monomer sequence, we can distinguish two basic types. These two types show up as more or less distinct peaks in the sequence distribution (Fig. 2a). On one hand, large, complex replicators with strong replicase activity and high fidelity exist, which also serve well as templates. On the other hand, there is a smaller peak of short molecules with limited function as replicases, but with high template quality. These act as parasites of the dominant quasispecies: they could not exist without the replicase but themselves lacking replicase activity they do not contribute to the 'common good'.

Local interactions and limited dispersion guarantee that the parasites are not able to spread beyond a certain limit; therefore they cannot competitively exclude good replicators. To demonstrate the key role spatial structure plays in restraining parasites, we constructed the mean-field counterpart of the model, in which each template is copied by an 'average replicase': a molecule possessing the average of replicase activity and fidelity over the whole grid. In the mean-field model, starting from an initial population of replicators built up in previous runs, the density of individuals decreases abruptly, and ultimately the whole population goes extinct. The spatial model, incorporating diffusion¹⁷, reveals the causes of this process. Increasing diffusion destroys the clustering of the altruists; replicator molecules gradually lose fidelity; all three replicator functions decline, template quality declines last, (Fig. 3). We recorded the ratio of 'birth' and movement events corresponding to particular diffusion rates. A slight diffusion does

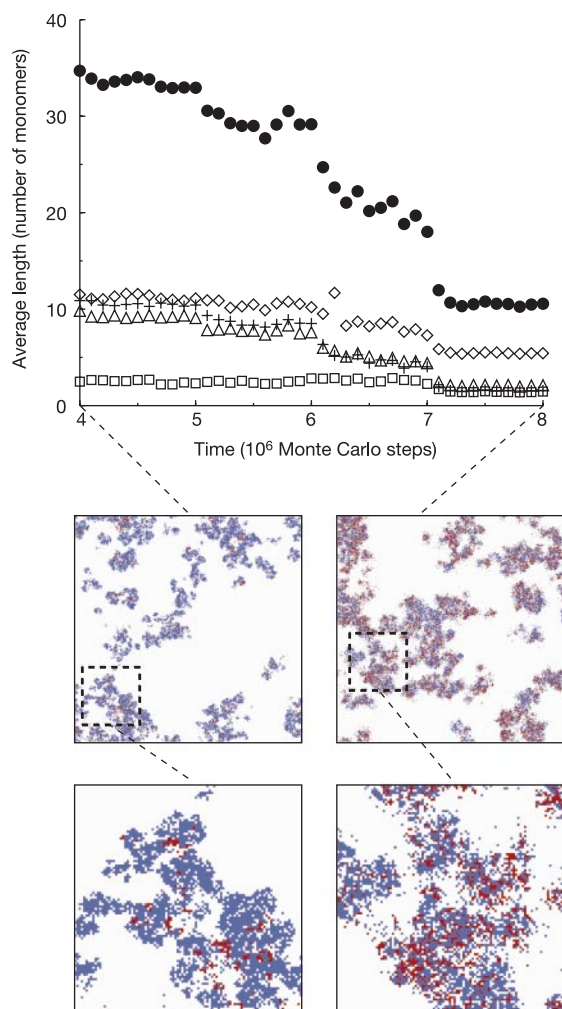


Figure 3 Replicator model with diffusion (population-wide averages). We start from the final state of Fig. 1 and increase diffusion rate step by step. Parameters as in Fig. 1, except diffusion rate: At $t = 4,000,000$ $D = 0.00001$ (replication event/movement due to a diffusion ratio of 12.5) and D jumps to 0.0001 (replication event/movement due to diffusion ratio of 1.6); 0.0002 and 0.0003 after every 10^6 Monte Carlo steps. Plates show the spatial distribution of replicators at $t = 4,000,000$ and $t = 8,000,000$ on two spatial scales. Upper plates show the whole arena; lower plates display an enlarged 100×100 part of it. Parasite and altruist replicators are distinguished on the basis of their lengths, according to Fig. 2a, b. Parasites (replicators, shorter than 25) are represented in red; altruist replicators (at least 25 monomers) are represented in blue. Empty sites are white.

not alter the results, and if the rate of birth and movement events is of the same order of magnitude, the results remain qualitatively the same. Impeding the evolution requires substantial mixing (see Fig. 3 legend). Our results agree with those from previous spatially explicit population dynamics of replicators, assuming either surface bonding^{18,19} or protocellular compartmentation²⁰.

To show the robustness of the results we have done additional runs with different parameter values in the functions $t(n_A)$, $r(n_B)$. This difference between the functions is expressed by their half-saturation values (v_h). The results did not show qualitative differences: in all cases the same evolutionary process takes place. At higher v_h values longer sequences build up, because in this case the assemblage of effective catalytic sites requires more monomers.

We consider what would happen to the system using an analogous algorithm assuming complementary base pairing. Owing to the tradeoffs involved, it is quite conceivable that a strand complementary to an excellent replicase would be an excellent template. This may aid the spread of function and would point into the spontaneous emergence of complementary genic and enzymatic strands, and a primitive form of transcription (biased replication), in the spirit of the qualitative scenario of ref. 21.

The aim of our model is to show that certain molecular traits, crucial to the origin of life, can spread under plausible circumstances. We note that the selective scenario simulated here is compatible with existing theories and experiments. Examples of such approaches are the 'polymerization on the rocks'^{22,23} (that is, chemical evolution on mineral surfaces likened to a primordial 'crêpes'²⁴) and the surface-promoted amplification^{25,26} chemical scenarios. Moreover, our simulation strengthens the view^{18,19} that selection dynamics on mineral surfaces could have had an important positive effect on the dynamical coexistence of useful molecules of a primitive genetic system. How exactly, from a chemical point of view, this important phase of prebiotic evolution happened, is an open question. One possible scenario is based on ligation, leading from small templates and chemical ligation of long complementary building blocks, to long templates and enzymatic ligation of short substrates, culminating in template-directed polymerization by successive monomer addition⁴. The assumption used here that even very short oligonucleotides could act as rudimentary replicases is unrealistic¹. A recently selected ribozyme is able to catalyse polymerization on an external template, up to the length of 14 nucleotides with an average accuracy of 0.967 per nucleotide²⁷. There are three snags: the ribozyme itself is over 180 nucleotides long; its efficiency (speed) is not high enough; and template and copy are not separated. The basic message of the model remains valid if we shift the minimum replicase length to realistic values, but computation becomes more cumbersome. At present, the only hope is that an efficient non-enzymatic RNA replication system will be found^{1,2} that could produce molecules long enough for a generalized replicase function. Once this is shown, further evolution is feasible by reciprocal molecular altruism 'on the rocks'. □

Methods

We have simulated the evolution of a large initial population of short replicators in a stochastic cellular automaton, because we suspect that spatial aspects play a decisive role in the dynamics of this system. The replication surface is a 400×400 square grid of toroidal topology to avoid edge effects. Each cell of the grid may be empty, or occupied by a replicator. An occupied cell is characterized by the sequence of the replicator it harbours. In each Monte Carlo step every site of the grid is updated exactly once, in a random order, according to the following rules:

- (1) If the site is empty, nothing happens;
- (2) If it is occupied, then with probability P_d the resident replicator decays;
- (3) If the replicator survives, then it is a potential template and replicates (a new replicator is 'born') with probability $P_r = \frac{2}{3} d r_n^L$ (where s is the number of potential replicators on the four neighbouring sites; r is replicase activity of one of the neighbouring

replicators, chosen randomly; t is the template efficiency of the focal replicator; n is the number of monomers in the template.) d is a global variable that describes the density of available monomers for replicators. Its value is the same all over the grid, which is a reasonable assumption, if we assume that diffusion of monomers is fast. $d = \frac{1}{\mu} \sum_{i,j=1}^L n_{i,j} / L^2$, summed over the grid (μ is a monomer abundance determining constant, $n_{i,j}$ is the length of replicator at position i,j and L is the length of sides of the square grid).

In the diffusive version a diffusion step occurs, by using the Toffoli–Margolus algorithm on a randomly chosen 2×2 square with probability D .

It would be tempting to identify our abstract monomers with real RNA bases. For example, in an experimental investigation of nonenzymatic template oligomerization⁹. It was found that cytidine acts like our monomer A. Yet our usage of the abstract monomers amounts to a subtle encoding as follows. Imagine that we assign a random (arbitrary) base sequence to a catalytic domain, such as AGGUGCCGAA. Then we encode this domain for our simulation as, say, BBBB BBBB. So abstract 'B' means any real base improving the activity in the given domain. A single base change may either destroy any function (which we handle by the dummy D), improve both (which is then no problem for evolution), or improve one function and diminish another.

The algorithm including diffusion is available at <ftp://hera.colbud.hu/users/szathmary/replik.c>

Received 8 May; accepted 13 September 2002; doi:10.1038/nature01187.

1. Joyce, G. F. & Orgel, L. E. in *The RNA World*, 2nd edn (eds Gesteland, R. F., Cech, T. R. & Atkins, J. E.) 49–77 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1999).
2. Eigen, M. Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465–523 (1971).
3. Poole, A., Jeffares, D. & Penny, D. Early evolution: the new kids on the block. *BioEssays* **21**, 880–889 (1999).
4. James, K. D. & Ellington, A. D. The fidelity of template-directed oligonucleotide ligation and the inevitability of polymerase function. *Orig. Life Evol. Biosphere* **29**, 375–390 (1998).
5. Scheuring, I. Avoiding catch-22 of early evolution by stepwise increase in copying fidelity. *Selection* **1**, 135–145 (2000).
6. Wolfram, S. *A New Kind of Science* (Wolfram Media, Champaign, Illinois, 2002).
7. Trivers, R. L. The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57 (1971).
8. Nowak, M. & Sigmund, K. Tit for tat in heterogeneous populations. *Nature* **355**, 250–253 (1992).
9. Inoue, T. & Orgel, L. E. A nonenzymatic RNA polymerase model. *Science* **219**, 859–862 (1983).
10. Florentz, C. & Giegé, R. tRNA: *Structure, Biosynthesis, Function* (eds Söll, D. & Rajbhandary, U. L.) 141–164 (American Society of Microbiology Press, Washington, 1995).
11. Galas, D. J. & Branscomb, E. W. Enzymatic determinants of DNA polymerase accuracy. Theory of coliphage T4 polymerase mechanisms. *J. Molec. Biol.* **124**, 653–687 (1978).
12. Goodman, M. F. & Fyngson, D. K. DNA polymerase fidelity: From genetics toward a biochemical understanding. *Genetics* **148**, 1475–1482 (1998).
13. Wang, J. *et al.* Crystal structure of a pol α family replication DNA polymerase from bacteriophage RB69. *Cell* **89**, 1087–1099 (1997).
14. Doublé, S. *et al.* Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**, 251–258 (1998).
15. Kiefer, J. R. *et al.* Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature* **391**, 304–307 (1998).
16. Steitz, T. A. Structural biology: A mechanism for all polymerases. *Nature* **391**, 231–232 (1998).
17. Toffoli, T. & Margolus, N. *Cellular Automata Machines: A New Environment for Modeling* 155–167 (MIT Press, Cambridge, Massachusetts, 1987).
18. Boerlijst, M. C. & Hogeweg, P. Spiral wave structure in prebiotic evolution: Hypercycles stable against parasites. *Physica D* **48**, 17–28 (1991).
19. Czárán, T. & Szathmáry, E. *The Geometry of Ecological Interactions: Simplifying Spatial Complexity* 116–134 (Cambridge Univ. Press, Cambridge, 2000).
20. Szathmáry, E. & Demeter, L. Group selection of early replicators and the origin of life. *J. Theoret. Biol.* **128**, 463–486 (1987).
21. Szathmáry, E. & Maynard Smith, J. The origin of chromosomes II. Molecular mechanisms. *J. Theoret. Biol.* **164**, 447–454 (1993).
22. Ferris, J. P. *et al.* Synthesis of long prebiotic oligomers on mineral surfaces. *Nature* **381**, 59–61 (1996).
23. Orgel, L. E. Polymerisation on the rocks: Theoretical introduction. *Orig. Life Evol. Biosphere* **28**, 227–234 (1998).
24. von Kiedrowski, G. Primordial soup or crêpes? *Nature* **381**, 20–21 (1996).
25. von Kiedrowski, G. Surface-promoted replication and exponential amplification of DNA analogues. *Nature* **396**, 245–248 (1998).
26. von Kiedrowski, G. & Szathmáry, E. Selection versus coexistence of parabolic replicators spreading on surfaces. *Selection* **1**, 173–179 (2000).
27. Johnston, W. K. *et al.* RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* **292**, 1319–1325 (2001).

Acknowledgements Financial support by the National Research Fund (OTKA) is gratefully acknowledged.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to E.S. (e-mail: szathmary@colbud.hu).

New dimensions in electrophoresis

Means of detection, data analysis and gel imaging.

2D tracking dye

Amresco www.amresco-inc.com

One to pursue

Amresco's 2D tracking dye is intended for use in two-dimensional polyacrylamide gel electrophoresis, in particular the company's 2D RITE-Gels. The dye accurately follows the progress of the second-dimension gel during electrophoresis. It is supplied pre-mixed and ready to use.

Zetalif

ESA Analytical www.esainc.com

Detecting not a lot...

The Zetalif laser-induced fluorescence detector monitors HPLC and capillary electrophoresis eluants down to zeptomoles for some compounds. That's sextillionths, or $\times 10^{-21}$ M. Zetalif uses monochromatic laser sources to detect trace levels of compounds with sensitivity claimed to be typically 100–1,000 times that of traditional fluorescence detection. The 14 operational wavelengths range from 266 to 785 nm and measurement is over a linear dynamic range of 3–4 orders of magnitude. The detector can connect to most HPLC, micro-HPLC, capillary LC, capillary electrophoresis and CEC systems.

AccuFLOWGel

Flowgen www.flowgen.co.uk

Band aids for electrophoresis

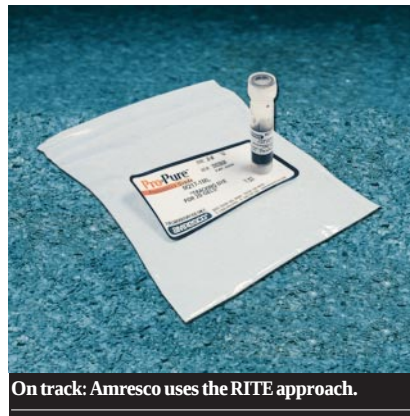
AccuFLOWGels are pre-mixed solutions, available in 30% or 40% concentrations, for protein and nucleic acid electrophoresis. AccuFLOWGel 19:1 is best suited for separating fragments below 100 bp. It contains 19 g acrylamide per g methylene bisacrylamide (MBA) cross-linker and provides a base solution for gels for automated sequencers. Urea can be added to the solution immediately prior to use. AccuFLOWGel 29:1 contains 29 g acrylamide per g MBA and is suitable for separating fragments of 100–1,000 bp. Applications include denaturing PAGE-DNA/RNA, gel electrophoresis of PCR products, ribonuclease protection and DNA sequencing.

E-Gel 96

Invitrogen www.invitrogen.com/egels

Staggering speed

E-Gel 96 is a pre-cast, robot-compatible agarose electrophoresis system that automates high-throughput analysis of multiple DNA samples. It uses staggered-well technology to perform simultaneous analysis of 96 samples in 12 minutes.



On track: Amresco uses the RITE approach.

GeneDirectory

Syngene www.syngene.com

Look through any window

The GeneDirectory data storage and analysis program works in conjunction with Syngene's GeneTools. It can be used to compare thousands of bands on different gels. Data can be viewed on screen as familial gel tracks, with up to five gel tracks containing DNA from known samples in an upper window and the bands on the gel to be compared in a lower one. The software offers three ways to compare data, ranging from simple band comparison to co-dominance and dominance analysis for RFLP and AFLP studies, forensic work or gene mapping.

ReadyStrip

Bio-Rad www.bio-rad.com

Go to great lengths

Bio-Rad has added an 18-cm immobilized pH gradient strip to its ReadyStrip product line for the first dimension of two-dimensional gel electrophoresis. The strips are permanently labelled to mark the anode and the pH range. ReadyStrips are available in broad, narrow and micro ranges, as well as a 6.3–8.3 range, in lengths of 7, 11, 17 and 18 cm. Applications include studying a well-characterized protein, mapping the proteome and isolating a large quantity of protein.

Model 5005 Mini Chiller

PolyScience www.polyscience.com

A cool response

The model 5005 mini-chiller from PolyScience combines the compact size of a circulator with an extended temperature range of -5°C to 50°C . Small enough to be placed on the bench top or under the bench, the 5005 offers up to 450 watts of cooling at 20°C . A

sealed plumbing system allows use in closed or open loop applications, and a microprocessor delivers the correct amount of refrigerant to handle the incoming heat load. The chiller is suitable for electrophoresis cell cooling, small bioreactors and general laboratory processes.

200 Series gas chromatograph

Cambridge Scientific Instruments

www.camsci.co.uk

Modern times

The 200 Series gas chromatograph is billed as an inexpensive way of replacing ageing gas chromatography systems. It provides the facilities and performance needed for most QC/QA or teaching applications, with minimal bulk and power consumption. A temperature-controlled, capillary gas chromatograph with split/splitless injector, FID and electronic gas control, the 200 Series GC has a footprint of 41×34 cm, weighs 7 kg and uses 700 watts of power.

Hydraplus autosampling system

Analytical Technologies

www.analyticaltechnologies.co.uk

Special delivery

The upgraded Hydraplus autosampling system is designed to automate the Hydrasys multiparameter electrophoresis system in clinical laboratories. The Hydraplus delivers precise micro-volumes of sample directly from primary tubes to four sample applicators in a built-in humidity chamber. Dilution ratios of actual concentrations can be pre-programmed and samples can be tracked from primary tube to sample applicator and gel through to the final result.

These notes are compiled in the Nature office from information provided by the manufacturers.

Contacts

Publisher: Ben Crowe

Editor: Paul Smaglik

Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Senior European Sales Manager:
Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)

Andy Douglas (4975)

Frank Phelan (4944)

Netherlands/ Italy/ Iberia:

Evelina Rubio Hakansson (4973)

Scandinavia: Silke Opstrup (4994)

France/ Belgium:

Amelie Pequignot (4974)

Production Manager: Billie Franklin

To send materials use London
address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

International

Advertising Coordinator:

Hind Berrada (4935)

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Ben Lund

European Satellite Office

Germany/ Austria/ Switzerland:

Patrick Phelan, Odo Wulffen

Tel + 49 89 54 90 57 11/-2

Fax + 49 89 54 90 57 20

e-mail: p.phelan@nature.com

o.wulffen@nature.com

US Head Office, New York

345 Park Avenue South,
10th Floor, New York, NY 10010-1707

Tel +1 800 989 7718

Fax +1 800 989 7103

e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

US Advertising Coordinator:

Linda Adam

Japan Head Office, Tokyo

MG Ichigaya Building (5F),

19-1 Harakutamachi,

Shinjuku-ku,

Tokyo 162-0841

Tel +81 3 3267 8751

Fax +81 3 3267 8746

e-mail: k.johnson@naturejp.com

Asia-Pacific Advertising Manager:

Keyn Johnson

naturejobs

Where are all the students?

Last week saw a two-day round-table meeting on government, university and industrial research take place at the National Academy of Sciences in Washington. At the end of the session, Marye Anne Fox, chancellor of North Carolina State University, spent 22 minutes summarizing both the problems and the possible solutions facing the US science and engineering workforce. One phrase she used repeatedly neatly encapsulates both — effort versus reward.

Fox noted that the number of US-born students entering graduate programmes in science and engineering has been slipping steadily for years. That hadn't been much of a problem until recently, because scientists from other countries readily filled the gap. But now there are signs that more foreign-born scientists trained in the United States are returning home.

Why the impending talent deficit? Fox's phrase may capture the reason. Over the past several years the cost-benefit formula for embarking on a scientific career has been tilting towards effort and away from reward. For example, the time it takes to earn a PhD has grown, as have the debts likely to be incurred.

Fox suggested several strategies to reverse those trends. Finding ways to speed up the PhD process and forgive loans for key areas of science and technology where shortages have been identified would be effective — especially for women and minorities who are disproportionately represented in science. Improving teaching and career advice at both the high-school and university levels would also help.

All of these ideas, of course, involve money. But even more difficult to secure are the required cultural changes. Promoting professors who do a great job teaching and mentoring? That may require too much effort and not enough reward for universities.

Paul Smaglik

Naturejobs editor



Contents

REGIONS

Life sciences get a boost
in Michigan

p4

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

SCIENTIFIC ANNOUNCEMENTS

SCIENTIFIC EVENTS



Corridor at the crossroads

Michigan

Charles Bisgaier believes a merger between Pfizer (above) and Pharmacia could inject fresh life into Michigan.



What happens when two roads that were running in parallel suddenly intersect? In Michigan, those employed in the life sciences may be about to find out. On the state's scientific-employment map, one track is the drug industry, represented by Pfizer in Ann Arbor and Pharmacia in Kalamazoo. The other road is a 20-year, \$50-million-a-year initiative known as the Michigan life-sciences corridor, designed to boost the state's standing in life sciences. Traffic along the two was supposed to drive the state from number ten in life-science employment into the top five. But with Pfizer shareholders set to vote on buying Pharmacia next month, that goal is in jeopardy.

The impending merger could take job traffic in two different directions. Redundancies could lead to a scientific exodus if there aren't enough institutions to absorb them. Pfizer has said that it wants to save \$2.5 billion after the merger, so it seems unlikely that all 2,800 positions at its Ann Arbor location and all 6,300 at Pharmacia's in Kalamazoo are safe. Or it could result in a renaissance, with some talent moving on to form spin-offs and others reinvigorating existing companies and academic research institutions.

Some signs point to the possibility of a brain drain. Indianapolis-based Eli Lilly has leased four billboards near the Pharmacia premises, with messages such as "Life at Lilly" and "Real people doing extraordinary things", accompanied by the company's web address.

Other, less literal signs are more sanguine. For instance, the earlier merger of Pharmacia and Upjohn gave rise to Esperion Therapeutics, when the new entity decided not to pursue a particular cardiovascular drug. And the company's management ranks are staffed by scientists from when Warner-Lambert and Parke Davis merged into Pfizer.

Esperion's co-founder and vice-president Charles

Bisgaier notes that the new merger is likely to create similar licensing opportunities. Money from the life-sciences corridor, which is being shared among basic and applied science, infrastructure and commercialization, may be "enough to prime the pump" when venture capital is tight, Bisgaier says.

The corridor, administered by the Michigan

Demand and supply

Michigan has become increasingly efficient at finding out what sorts of skills the area's life-science employers want in prospective employees, and then delivering them. This month, area university administrators met with company recruiters to discuss this issue.

Such input can be used to shape programmes such as Wayne State University's master's degree in biotechnology. That programme — involving a year of coursework, a year of lab courses and a summer placement in a local company — measures students' progress against regular

milestones, just as biotech companies are evaluated by venture capitalists. The programme attracts a few hundred applicants a year but can only take about 18 at a time. One of the reasons for its popularity, says the programme's director Phil Cunningham, may be that students who want to go into biotech may need graduate training, but not necessarily a PhD and a postdoc. And companies that need skilled people to perform research will not necessarily need a cadre of PhDs.

P.S.

Wayne State Molecular Biotech Master's Programme
bio.wayne.edu/biotech

Economic Development Council, is funded by a settlement between tobacco companies and US states to compensate for the medical costs of tobacco-related illness. It is already investing in some relevant start-up firms. In the latest round of funding, Velcura Therapeutics got \$3.3 million for research that should benefit cigarette smokers, whose habit reduces their ability to absorb the calcium needed to stop their bones becoming brittle. Velcura was spun off from founder Michael Long's research in growing bone outside the body. The corridor cash helped Long to move from his University of Michigan (UM) Medical School lab to a larger space off campus and to recruit more staff.

CORRIDOR OF POWER

The first round of funding, which encompassed 2000 and 2001 and administered about \$100 million, planted similar seeds. A commercialization grant of \$790,000 for ApoLife helped the Detroit start-up to get into the recombinant-protein business. In 2001, 22 companies started in Michigan, many with help from the corridor.

The corridor's core, though, will benefit both public and private research. The biggest investment is about \$70 million for facilities at the institutions that mark the corridor's boundaries. Michigan State University (MSU) at East Lansing houses a structural-biology centre, the UM at Ann Arbor hosts proteomics and

Private funds promise a positive outlook for the Van Andel Research Institute.

bioinformatics facilities, the Van Andel Institute in Grand Rapids contains an animal-model facility, and Wayne State University in Detroit acts as a genomics centre.

The distributed arrangement is designed to avoid duplication and to encourage sharing of resources. It will also benefit established companies such as Lumigen, which spun out of Wayne State in 1987. Paul Schaap, founder of the chemiluminescence-reagent producer, expects the company to use corridor resources such as high-

frequency NMR instrumentation.

Such bridge-building hasn't been limited to infrastructure. Although Michigan ranks among the top five states in terms of federal research funding, "there wasn't much collaboration" until a few years ago, says Raili Kerppola, managing director of the Michigan Life Sciences Corridor. "Universities didn't collaborate much. No one did."

One of the corridor's biggest grants to a researcher in 2000 — \$2.1 million over three years — allowed Jack Harkema, a veterinary pathologist and inhalation toxicologist at MSU, to form a partnership with Gerald Keeler's lab in the UM School of Public Health.

Harkema's project used a mobile research lab to collect air from various sites, filter and concentrate it, and then see how it affected animal models for human respiratory diseases such as asthma. The relationship between the two labs was complementary, with Harkema's group concentrating on the biology and Keeler's team sorting out the atmospheric science.

New facilities not funded by the corridor are poised for similar success. The \$60-million Van Andel Research Institute, funded by Amway founders Jay and Betty Van Andel, opened in Grand Rapids at about the same time as the corridor came into existence. At that time, there was little scientific activity in Grand Rapids. As George Vande Woude, the institute's director, says: "None of the principal investigators had ever heard of Grand Rapids." But once the building was erected and the recruits convinced that the institute would eventually be supported by a billion-dollar-plus endowment, the task of attracting researchers became easier. Hosting the transgenic-animal core facility plus a proteomics and genomics infrastructure also helps to attract investigators and postdocs. Once the building fills up, the institute will raise another, about 50% bigger, on land adjacent to the existing site.

Alan Saltiel, director of the UM Life Sciences Institute, hopes that the institute's \$100-million facility will also act as a talent magnet when it opens in September 2003. Saltiel, who left Parke Davis when the company was taken over by Pfizer in 2001, calls the project "exciting", because he plans to recruit biologists, chemists and physicists, and encourage them to interact as much as possible. The institute now has seven faculty members, including Saltiel, but will grow to 20–30 principal investigators once the institute, which is designed to hold 300–350 people, opens.

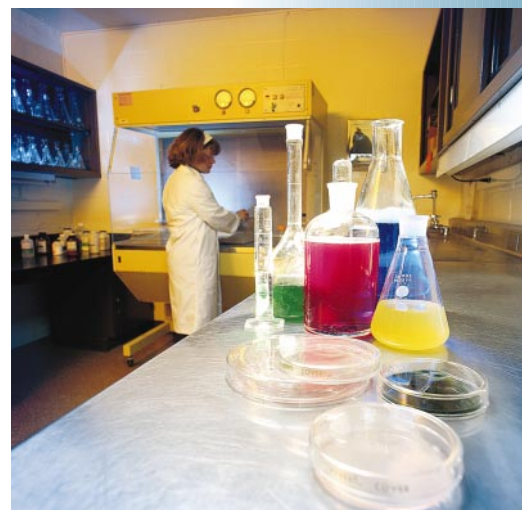
Saltiel is happy that, while at Parke Davis, he also held an adjunct position at the UM that allowed him to publish. His decision to stay in Michigan now seems sound. He is sharing in one of the corridor's biggest grants of 2002 — a \$3.5-million diabetes-research consortium that teams him up with some of his old colleagues now at Wayne State, as well as researchers from MSU. He suspects that the life-sciences corridor will provide an equally agreeable destination for scientists from Pfizer or Pharmacia.

Paul Smaglik is editor of *Naturejobs*.

Life Sciences Corridor medc.michigan.org/lifescience



Alan Saltiel is pleased that he has stayed in Michigan.



Boosted: with spin-offs and an influx of cash, life-science research in Michigan is robust.

Close call

In a referendum this month, Michigan voters defeated a proposal that would have gutted the area's Life Sciences Corridor just as the project was gaining momentum. The state now receives \$50 million a year from settlements with the tobacco industry. The proposal would have reduced that

amount by up to \$20 million a year, instead directing more money into healthcare and anti-tobacco advertising. Bob Huggett, vice-president of research at Michigan State University in East Lansing, says that the proposal would have "greatly impacted" research in the state, had it been accepted. P.S.